



# Elastodynamic Analysis for Slow Tectonic Loading with Spontaneous Rupture Episodes on Faults with Rate- and State-dependent Friction

## Citation

Lapusta, N., James R. Rice, Yehuda Ben-Zion, and Gutuan Zheng. 2000. Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research* 105(B10): 23,765–23,789

## Published Version

<http://dx.doi.org/10.1029/2000JB900250>

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:2668694>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

LAPUSTA ET AL.: DYNAMIC ANALYSIS OF SLOWLY LOADED FAULTS

To appear in the *Journal of Geophysical Research*, 2000.

## **Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction**

Nadia Lapusta and James R. Rice

Division of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts

Yehuda Ben-Zion

Department of Earth Sciences, University of Southern California, Los Angeles, California

Gutuan Zheng

Technology Group, IBM, Somers, New York

**Abstract.** We present an efficient and rigorous numerical procedure for calculating the elastodynamic response of a fault subjected to slow tectonic loading processes of long duration within which there are episodes of rapid earthquake failure. This is done for a general class of rate- and state-dependent friction laws with positive direct velocity effect. The algorithm allows us to treat accurately, within a single computational procedure, loading intervals of thousands of years and to calculate, for each earthquake episode, initially aseismic accelerating slip prior to dynamic rupture, the rupture propagation itself, rapid post seismic deformation which follows, and also ongoing creep slippage throughout the loading period in velocity-strengthening fault regions. The methodology is presented using the two-dimensional (2-D) antiplane spectral formulation and can be readily extended to the 2-D in-plane and 3-D spectral formulations and, with certain modifications, to the space-time boundary integral formulations as well as to their discretized development using finite difference or finite element methods. The methodology can be used to address a number of important issues, such as fault operation under low overall stress, interaction of dynamic rupture propagation with pore pressure development, patterns of rupture propagation in events nucleated naturally as a part of a sequence, the earthquake nucleation process, earthquake sequences on faults with heterogeneous frictional properties and/or normal stress, and others. The procedure is illustrated for a 2-D crustal strike-slip fault model with depth-variable properties. For lower values of the state-evolution distance of the friction law, small events appear. The nucleation phases of the small and large events are very similar, suggesting that the size of an event is determined by the conditions on the fault segments the event is propagating into rather than by the nucleation process itself. We demonstrate the importance of incorporating slow tectonic loading with elastodynamics by evaluating two simplified approaches, one with the slow tectonic loading but no wave effects and the other with all dynamic effects included but much higher loading rate.

## 1. Introduction

The purpose of this paper is to establish an efficient algorithm for elastodynamic shear rupture analysis of a fault governed by a general class of rate- and state-dependent friction laws in situations for which the total time of loading is vastly longer than the time for waves to traverse the domain of interest. Such an algorithm is needed to study slow tectonic loading processes during which there are episodes of spontaneous rapid failure in earthquakes. Investigating these processes requires a special approach, since quasi-static methods (used for calculating slow deformational processes of long duration) fail as instabilities develop, while standard elastodynamic algorithms not only use relatively small time steps but also require an increasing amount of memory and computational time at each time step to take into account all the prior deformation history, and hence they are excluded from direct implementation for investigating long-duration processes because of limitations on computing resources.

Various solutions have been proposed. One of them [e.g., Okubo, 1989; Shibazaki and Matsu'ura, 1992] is to employ a quasi-static method during slow deformation and then to switch to a dynamic method once an instability starts. However, the abrupt switching from one scheme to another may disrupt the natural development of the instability, and the effects of this disruption on the further model response cannot be easily determined within this approach. Other approaches [e.g., Cochard and Madariaga, 1996; Myers *et al.*, 1996] neglect all aseismic fault slippage, so that stressing between earthquakes is trivially modeled, and give the fault a "kick" in the form of an abrupt small strength drop, once a critical stress has been reached somewhere. At that stage an elastodynamic algorithm calculates rupture until arrest occurs. This inevitably generates a population of small ruptures, and it requires careful study of dependence on the abrupt strength drop magnitude to separate which may be physical and which are artifacts of the abrupt drop [Cochard and Madariaga, 1996]. Still another alternative is to use a plate loading rate which is only a few orders of magnitude less than representative seismic slip rates, rather than the roughly 10 orders as for natural faults, and to use standard elastodynamic numerical methodology throughout (like in the work by Shaw and Rice [2000]). This is straightforward to implement, at least if some provision is made for dissipating wave energy, but makes it difficult to suitably model aseismic slip processes and can blur the distinction between aseismic slip before instability and small earthquakes.

The developments of the present work provide an integrated numerical scheme allowing resolution of both slow and fast deformational phases, as well as the transition between them, within a single mathematical framework for

elastodynamics. The method enables us to perform calculations over thousands of years of slow tectonic loading, punctuated by earthquakes and the processes which lead to and follow them. Thus we can resolve aseismic slip on velocity-strengthening fault regions, advance of slip into more firmly locked zones, and slowly accelerating aseismic slippage that grows in spatial extent and will ultimately break out into an earthquake but has duration that is vastly longer than the seismic event itself. We also resolve all details of the break out of rupture, its propagation and arrest, and the transient post seismic slippage that develops.

Our methodology for studying slow loading processes has two main ingredients. The first is based on the form of elastodynamic relations that we use, in which the dependence of the inertial response on prior deformation history can be truncated so that only a (fixed) part of the deformation history back from current time needs to be considered. That translates into fixed memory requirements and fixed amount of computation per each time step. It also makes the computation at each time step independent of how much time has already been simulated. The methodology is illustrated in this paper for the two-dimensional (2-D) antiplane case and uses a spectral representation of elastodynamic relations developed by Perrin *et al.* [1995] in which the slip distribution is represented as a Fourier series in the spatial coordinate, truncated at large order, and fast Fourier transform (FFT) methods are used. The corresponding methodology for the 2-D in-plane and 3-D cases is conceptually very similar and can be easily adopted from the one presented here using 3-D spectral elastodynamic relations developed by Geubelle and Rice [1995] and Cochard and Rice [1997]. Our algorithm can also be generalized to the (closely related) space-time boundary integral formulation. Furthermore, for situations such as elastic property heterogeneity that are not congenial to spectral or boundary integral approaches, finite difference or finite element procedures could be used, not in their conventional application to directly calculate the rupture propagation itself, but rather to calculate and numerically tabulate the convolution kernels for use in our methodology.

The second ingredient is variable time stepping. The size of the time step to be made is dictated by the current values of slip velocities and parameters of the constitutive law. The smaller the slip velocities, the larger the time step, and vice versa. While the truncation of the convolutions over prior slip velocity history reduces the amount of computation required to complete one time step, the variable time stepping reduces enormously the number of time steps needed to simulate processes during the essentially aseismic phases of deformation which constitute almost all of the fault history. Throughout the computation, time steps can change by many orders of magnitude in value, allowing us to go in rel-

atively few steps through periods of essentially quasi-static loading, to consider more carefully the nucleation phase, and to resolve in great detail the features of the dynamic propagation during an instability. The coefficients of proportionality between the time steps and slip velocities depend on the parameters of the constitutive law as well as on numerical stability considerations that we derive here. We present the formulation for a general class of rate- and state-dependent friction laws with a positive direct velocity effect. The prototype of such laws is the experimentally derived logarithmic law of *Dieterich* [1979, 1981] and *Ruina* [1983]. The presence and size of the positive direct velocity effect for the quasi-static range of slip velocities, amply documented in such experiments, are shown to be crucial in allowing long time steps during slow deformation phases without losing stability (during such phases, velocity-strengthening parts of the fault zone are continuously slipping, producing an aseismic viscoplastic type response to which rate- and state-dependent friction then reduces).

The main goal of the present paper is to give the detailed description of the method in its current, much improved form. Earlier versions of the methodology were briefly outlined and/or implemented by *Zheng et al.* [1995], *Rice and Ben-Zion* [1996], and *Ben-Zion and Rice* [1997]. We describe the algorithm ingredients in sections 2-6. The new developments include understanding constraints on the time step during slow deformation phases and corresponding limitations on the procedure applicability (section 4), much more efficient truncation and evaluation of the convolution integrals involved (section 6), and a new procedure for updating the system in a time step (section 5). These developments allow consideration of a much wider range of the constitutive parameters, better numerical convergence of the results, and enhanced resolution in time and space with the same computational resources.

The proposed methodology can be used to address a number of important issues, such as fault operation under low overall stress, interaction of dynamic rupture propagation with pore pressure development, patterns of rupture propagation in events nucleated naturally as a part of a sequence, the earthquake nucleation process, earthquake sequences on faults with heterogeneous frictional properties and/or normal stress, and others. Section 7 demonstrates the implementation of the algorithm by considering the elastodynamic response of a 2-D crustal strike-slip model, with depth-variable properties, descended from the model of *Tse and Rice* [1986] and studied by *Rice and Ben-Zion* [1996] and *Ben-Zion and Rice* [1997]. Considering a wider range of constitutive parameters than the range tractable for previous studies, we observe that small events appear for lower values of the state-evolution distance. The nucleation phases of the

small and large events are very similar, suggesting that the size of an event is determined by the conditions on the fault segments that the event is propagating into rather than by the nucleation process itself. We show how insufficient resolution in time can produce more complex slip accumulation that looks "smooth" and plausible yet is just a numerical artifact. We also evaluate two simplified approaches, one with the slow tectonic loading but no wave effects (quasi-dynamic approach, as in the work by *Rice* [1993]), and the other with all dynamic effects included but much higher loading rate (like in the work by *Shaw and Rice* [2000]). The comparison shows that incorporating slow tectonic loading is very important for determining the true model response.

## 2. Elastodynamic Relation and Truncation of Convolution Integrals

As an illustration of the elastodynamic relations, let us consider a 2-D antiplane framework, in which the fault plane coincides with the  $x$ - $z$  plane of a Cartesian coordinate system  $xyz$  and all particles move parallel to the  $x$  direction. The only nonzero displacement is  $u_x(y, z, t)$ , and we define slip  $\delta(z, t)$  on the fault plane as the displacement discontinuity  $\delta(z, t) = u_x(0^+, z, t) - u_x(0^-, z, t)$ . The relevant shear stress on the fault plane is denoted by  $\tau(z, t) = \sigma_{yx}(0, z, t)$ . It is possible to express the stress on the fault plane in terms of the slip history on the fault plane only [e.g., *Cochard and Madariaga*, 1994, *Perrin et al.*, 1995] as

$$\tau(z, t) = \tau^o(z, t) + f(z, t) - \frac{\mu}{2c} V(z, t), \quad (1)$$

where  $\mu$  is the shear modulus,  $c$  is the shear wave speed,  $V(z, t) = \dot{\delta}(z, t) = \partial\delta(z, t)/\partial t$  is the slip rate,  $\tau^o(z, t)$  is the "loading" stress (i.e., the stress that would act if the fault plane  $y = 0$  were constrained against any slip), and  $f(z, t)$ , incorporating stress transfers, is a linear functional of prior slip  $\delta(z', t')$  over the causality cone (i.e., all  $z'$  and  $t'$  satisfying  $c(t - t') > |z - z'|$ ). The last term of (1) represents radiative damping [*Rice*, 1993], and the explicit extraction of the damping term from the functional  $f(z, t)$  allows for evaluation of  $f(z, t)$  without concern for singularities.

In (1), most of the elastodynamic response is contained in the stress transfer functional  $f(z, t)$ . *Cochard and Madariaga* [1994] have expressed it as a double convolution integral in space and time. *Perrin et al.* [1995] have derived a spectral representation of  $f(z, t)$  as a single convolution integral in time for each Fourier mode, when representing slip and the functional as Fourier series in space. Generalizations to general slip and/or opening states in 3-D problems are given by *Geubelle and Rice* [1995] and *Cochard and Rice* [1997]. We use the spectral representation of *Per-*

rin et al. [1995] for the illustration here and write

$$\delta(z, t) = \sum_{n=-N_{\text{ele}}/2}^{N_{\text{ele}}/2} D_n(t) e^{ik_n z}, \quad (2)$$

$$f(z, t) = \sum_{n=-N_{\text{ele}}/2}^{N_{\text{ele}}/2} F_n(t) e^{ik_n z}, \quad k_n = \frac{2\pi n}{\lambda},$$

where  $\lambda$  is the length of the fault domain under consideration, replicated periodically. The replication distance  $\lambda$  has to be chosen several times larger than the domain over which rapid faulting takes place, to assure that there is negligible influence of waves arriving from the periodic replicates of the rupture process. For adaptation to our numerical procedure,  $N_{\text{ele}}$  (even) will be some large number of FFT sample points used to discretize this domain. Also, coefficients  $D_n$  and  $F_n$  are complex in general, with respective conjugates  $D_{-n}$  and  $F_{-n}$ , but take real values for  $n = 0$  and  $N_{\text{ele}}/2$ , so that the representations involve  $N_{\text{ele}}$  degrees of freedom. To satisfy the elastodynamic wave equation,  $D_n(t)$  and  $F_n(t)$  are related by

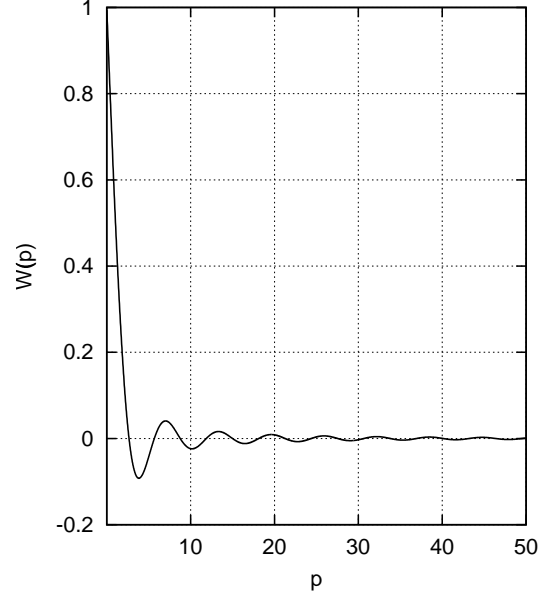
$$F_n(t) = -\frac{\mu |k_n|}{2} \int_0^t \frac{J_1(|k_n| ct')}{t'} D_n(t-t') dt', \quad (3)$$

where  $J_1()$  is the Bessel function of the first kind of order one. Equations (1)-(3) are referred to as the "displacement" representation of the elastodynamic relations. An analogous "velocity" representation can be obtained by combining (1) and (2) with the result of integrating (3) by parts, giving

$$F_n(t) = -\frac{\mu |k_n|}{2} D_n(t) + \frac{\mu |k_n|}{2} \int_0^t W(|k_n| ct') \dot{D}_n(t-t') dt', \quad (4)$$

where  $\dot{D}_n(t) = dD_n(t)/dt$  and  $W(p) = \int_p^\infty [J_1(\xi)/\xi] d\xi$  with  $W(0) = 1$ . As any other boundary integral formulation, the spectral representation (2), (3) or (2), (4) gives  $f(z, t)$  as a functional of  $\delta(z, t)$ , because  $f(z, t)$  can be expressed in terms of the  $F_n(t)$ ,  $F_n(t)$  are related to  $D_n(t)$ , and  $D_n(t)$  can be expressed in terms of  $\delta(z, t)$  by the inverse Fourier transform.

We emphasize that the spectral representation, in comparison with space-time boundary integral formulations, is very advantageous from the computational point of view. The matrix of convolution integrals, implied by a space-time formulation after discretization in space, is replaced in the spectral approach by a diagonal matrix, once the FFT



**Figure 1.** Convolution kernel  $W(p)$  for the velocity formulation in the 2-D antiplane case.

is used to transform from  $\delta(z, t)$  to  $D_n(t)$  and then the inverse FFT from  $F_n(t)$  to  $f(z, t)$ . Even though the spectral approach uses a larger number of degrees of freedom than needed for the domain of interest itself, the drastic reduction in the number of time convolutions significantly shortens the computation of the stress transfer functional  $f(z, t)$ , which is the most time-consuming stage of the analysis. We show this in Appendix B, where we further discuss the relation between spectral and space-time formulations. Note that *Cochard and Rice* [1997] showed how to reformulate the spectral method to rigorously eliminate the replications, but that requires far more complex calculations of the convolution kernels and still twice more degrees of freedom than needed for the domain of interest.

If the convolution integrals in (3) or (4) had to be computed in full, the algorithm would be impractical for investigation of long deformational processes. Evaluation of the convolution integrals is the most computationally demanding part of the elastodynamic analysis and may take more than 99% of the total computational time [*Perrin et al.*, 1995]. Fortunately, truncation of the convolutions is possible, which significantly reduces the overall computational time. If the duration of the physical problem is much longer than the time required for elastic waves to traverse the spatial domain of the system, it is not necessary to keep examining the influence of displacements of points on the failure surface at all prior times. This is reflected in rapid decay of

the kernels of the convolution integrals for both the velocity and displacement formulations. At large values of  $p$ , the kernels  $W(p)$  (shown in Figure 1) and  $J_1(p)/p$  have amplitude decay like  $1/p^{3/2}$  for an oscillation (at circular frequency  $|k_n|c$ ) that averages in time to zero.

To truncate the convolutions, we define the elastodynamic time window as the time interval  $[t - T_w, t]$ , where, in the computationally most efficient version of our methodology,  $T_w$  may be different for different Fourier modes.  $T_w$  is chosen such that the contribution to the functional  $f$  of the deformation history occurring prior to time  $(t - T_w)$  has negligible effect on the simulation results. Since only the effects from the current time  $t$  backward to  $(t - T_w)$  need be included in the dynamic response, the convolution integrals are truncated by computing them only within the elastodynamic time window defined. This transforms (3) and (4) into

$$F_n(t) = -\frac{\mu |k_n|}{2} \int_0^{T_w} \frac{J_1(|k_n|ct')}{t'} D_n(t-t') dt' \quad (5)$$

and

$$F_n(t) = -\frac{\mu |k_n|}{2} D_n(t) + \frac{\mu |k_n|}{2} \int_0^{T_w} W(|k_n|ct') \dot{D}_n(t-t') dt', \quad (6)$$

respectively. The discussion of the truncation implementation is given in section 6.

In view of the truncation procedure, the velocity formulation has an important advantage over the displacement formulation. As pointed out by Perrin *et al.* [1995], the first (algebraic) term in (4) and (6) corresponds to the final static elastic stress, and the remaining integral term corresponds to wave-mediated stress transfer carrying the elastodynamic effects. The isolation of the static term is important in our computational procedure where we truncate the remaining convolution integral. During slow deformational periods where  $\dot{D}_n(t)$  is small, the static term  $-\mu |k_n| D_n(t)/2$  contributes most to  $F_n(t)$ . During all deformation phases, the velocity formulation with truncation (1), (2), and (6), unlike the displacement formulation with truncation (1), (2), and (5), ensures that regardless of the way the convolution integral is truncated, the long-term static stress field (that is, the stress field after passage of all waves) due to slip up to the time  $t$  is always exactly represented. Thus, the velocity formulation should be used for long deformational histories, although the displacement formulation can also be useful, for example, to study individual events. The same separation into static and dynamic parts, with truncation of the convolution on time within the dynamic part, may be carried out in

the framework of the space-time representation for  $f(z, t)$ , as we briefly discuss in Appendix B.

Note that the combination of (1), (2), and (6) with  $T_w = 0$  (no convolution) would amount to static calculation of stress transfers, then corresponding to the "quasi-dynamic" procedure of Rice [1993], also discussed by Ben-Zion and Rice [1995] and Rice and Ben-Zion [1996]. Because of the retention of the radiation term of inertial elastodynamics, as  $\mu V/(2c)$  in (1), the quasi-dynamic procedure allows solutions to exist during instabilities; the solutions would not exist in a formulation with no damping term, which we usually call quasi-static.

### 3. Constitutive Laws and Space Discretization

Constitutive laws used here are rate- and state-dependent friction laws developed to incorporate experimental observations [Dieterich, 1979, 1981; Ruina, 1983]. These laws include dependence of strength on slip velocity and on an evolving state variable (or variables) which characterizes asperity contacts, thus allowing for loss of strength in rapid slip and for subsequent rehealing so that repetitive failures can occur. The laws have been successfully used to explain various aspects of stable and unstable sliding between elastic solids [Ruina, 1983; Rice and Ruina, 1983; Gu *et al.*, 1984; Tullis and Weeks, 1986] as observed in the laboratory. Also, they have been used to model earthquake phenomena, including nucleation, ductile and brittle crustal slip regions, spatio-temporal slip complexities, and earthquake aftershocks [e.g., Tse and Rice, 1986; Stuart, 1988; Okubo, 1989; Horowitz and Ruina, 1989; Rice, 1993; Dieterich, 1992, 1994; Perrin *et al.*, 1995; Ben-Zion and Rice, 1995, 1997; Rice and Ben-Zion, 1996; Stuart and Tullis, 1995; Tullis, 1996; Boatwright and Cocco, 1996].

A formulation of such laws which assumes constant normal stress and one state variable, to record dependence on slip history, is of the general form

$$\tau = \psi(V, \theta), \quad (7a)$$

$$d\theta/dt = \varphi(V, \theta), \quad (7b)$$

where  $\theta$  is the state variable and  $\tau$ ,  $V$ , and  $\theta$  depend on space variables and time. The rate- and state-dependent constitutive laws as usually formulated, based on laboratory observations, have the following properties. If slip velocity  $V$  is held constant, the state variable and hence the stress evolve toward constant values, called steady-state values and denoted  $\theta_{ss} = \theta_{ss}(V)$  and  $\tau_{ss} = \tau_{ss}(V)$ , respectively, where  $\theta_{ss}$  satisfies  $\varphi(V, \theta_{ss}) = 0$  and  $\tau_{ss}$  is given by  $\tau_{ss} = \psi(V, \theta_{ss})$ . All laws of the class (7) reduce, for  $\theta$  near  $\theta_{ss}(V)$ , to  $d\theta/dt \approx -(V/L)[\theta - \theta_{ss}(V)]$ , where  $L$  has dimensions of slip and can be interpreted as a characteristic

slip distance required for evolution to the steady state ( $L$  is also sometimes denoted by  $d_c$  or  $D_c$ ). The state variable  $\theta$  is usually chosen in such a way that  $\theta_{ss} = L/V$ , particularly if  $\theta$  is to be interpreted as a characteristic lifetime of the asperity population on the contact surfaces;  $L$  is then interpreted as a measure of the sliding distance required to establish a new population of asperity contacts, and is assumed to be independent of  $V$ . The law (7) is said to exhibit steady-state velocity weakening if  $d\tau_{ss}/dV < 0$  and steady-state velocity strengthening if  $d\tau_{ss}/dV > 0$ . If the slip velocity  $V$  is suddenly increased or decreased, the stress  $\tau$  simultaneously increases or decreases; that is, instantaneous positive viscosity is incorporated in (7) through the requirement

$$\frac{\partial\psi(V, \theta)}{\partial V} > 0. \quad (8)$$

This property is sometimes called "direct velocity dependence" or "direct effect" and is well established experimentally. As discussed in section 4, the presence and size of this direct effect are essential for our numerical procedure to be efficient in simulating processes of long duration. Hence our method is not immediately applicable to other constitutive laws, such as slip-weakening laws (which emerge as the limit case here for rapid slip if  $\psi$  and  $\varphi/V$  are, in the limit, independent of  $V$ ), which do not have that property.

Stability of steady frictional sliding, governed by the constitutive laws of type (7) with the properties discussed above, has been extensively investigated [Ruina, 1983; Rice and Ruina, 1983; Dieterich, 1992; Gu et al., 1984; Ranjith and Rice, 1999], particularly for single degree of freedom elastic systems. Such systems are generically represented by a spring-slider model, in which a rigid block is attached to a spring of stiffness  $k$  and slides on a frictional surface, with the other end of the spring moving at the imposed rate  $V^*$ . Linear stability analysis of such a system, perturbed about steady-state sliding at the rate  $V^*$ , as in the work by Ruina [1983], shows that the sliding is always stable for friction with steady-state velocity strengthening, while for friction with steady-state velocity weakening, there exists a critical value of the spring stiffness  $k_{cr}$  such that perturbations from steady-state sliding grow in time for systems with  $k < k_{cr}$  and decay in time for systems with  $k > k_{cr}$ . For rates  $V^*$  sufficiently small so that the inertia effects can be ignored, the critical stiffness is given by

$$k_{cr} = \left[ -\frac{V d\tau_{ss}(V)/dV}{L} \right]_{V=V^*}, \quad (9a)$$

where the precise definition of  $L$  is

$$L = [-V/(\partial\varphi(V, \theta)/\partial\theta)]_{V=V^*, \theta=\theta_{ss}(V^*)}. \quad (9b)$$

Here and in the following, notation  $[ \ ]_{V=V^*, \theta=\theta_{ss}(V^*)}$  means that the expression in the brackets has to be evaluated

at steady state given by  $V = V^*$ ,  $\theta = \theta_{ss}(V^*)$ . The origin of the well-known result (9a) is important for our consideration of variable time stepping in section 4. We review it in Appendix A and use it to restrict the size of time steps.

Such a stability dependence on the system stiffness in the case of steady-state velocity weakening has important implications for the proper space discretization, imposing an upper bound on a spatial element size in numerical modeling. To demonstrate this, let us continue with the antiplane example. Selecting  $N_{ele}$  equally spaced sample points along the domain of length  $\lambda$ , we discretize the domain into space elements (also called "cells")  $[z_{i-1}, z_i]$ ,  $z_i = ih$ ,  $i = 1, 2, \dots, N_{ele}$ ,  $h = \lambda/N_{ele}$ . The discretized formulation deals with slips  $\delta_i$  and shear stresses  $\tau_i$  at the sample points, taken at the cell centers ( $z_i - h/2$ ). Each of the cells has the effective stiffness  $k$  (defined as reduction in  $\tau_i$  due to elastic interactions with the surroundings for unit slip  $\delta_i$  at the same sample point), given by  $k = \gamma\mu/h$ . Here  $\gamma$  is a model-dependent constant, of order unity. For example,  $\gamma = 2/\pi$  when using the cellular basis set for slip (i.e., calculating the  $\tau_i$  as if the slip were locally uniform in each cell) like in the work by Rice [1993] and  $\gamma = \pi/4$  for the spectral basis set of (2). The values of  $\gamma$  cited apply for cells whose distance from any free surface is many times  $h$  and, in the spectral case, for  $N_{ele} \gg 1$ .

If we start with steady quasi-static sliding of the whole domain at slip rate  $V^*$  (taking this rate to be sufficiently small so that the dynamic effects are negligible) and slightly perturb the motion of one cell while maintaining steady sliding of the other cells, we get that the linearized response to the perturbation is governed by the same system of equations as for the spring-slider model, with the spring stiffness replaced by the effective stiffness of the cell. Hence the perturbation will grow if  $k = \gamma\mu/h < k_{cr}$  or, equivalently,  $h > \gamma\mu/k_{cr}$ , and the perturbation will decay if  $k = \gamma\mu/h > k_{cr}$  or, equivalently,  $h < \gamma\mu/k_{cr}$ . This property defines the critical cell size  $h^* = \gamma\mu/k_{cr}$ . As emphasized by Rice [1993], the growth of the perturbation on one cell, while the others continue the steady sliding, would imply that the cell is capable of failing independently of the surrounding cells, which would make the results dependent on the numerical discretization. Hence, to insure that the perturbation on a single cell decays, so that each cell can fail only as a part of larger space segment, the mesh should be refined enough for space element size  $h$  to be much smaller than the critical cell size  $h^*$ . In other words, the condition

$$h^*/h \equiv \eta \gg 1, \quad (10)$$

$$h^* = \gamma\mu/k_{cr} = -\frac{\gamma\mu L}{\max[V d\tau_{ss}(V)/dV]}$$

should hold, where the maximum is sought over all allowable slip rates  $V$ .

The critical cell size  $h^*$  is directly related to the nucleation size of model earthquakes, that is, to the size of the patch that initiates unstable, dynamic slip. In simulations with the constitutive laws of the type discussed here, rapid, dynamic break-out of an instability is always preceded by quasi-static slipping of a small zone. The size of this zone just before the dynamic instability is what we call the nucleation size. Changing  $L$ , and consequently  $h^*$ , changes the nucleation size in an essentially linearly proportional manner in all simulations we have done. Thus  $h^*$  is not only a very important numerical parameter, it is also a crucial physical parameter.

Note that expression (10) is derived neglecting inertial effects. *Rice and Ruina* [1983] have shown (by considering the spring-slider model with mass) that inclusion of the inertial effects increases  $k_{cr}$  in (9a) and hence decreases  $h^*$  in (10), which requires making cell size  $h$  even smaller. They have also analyzed uniform slip (at constant slip velocity) between elastic continua, with spatial perturbation of the type  $\sin(2\pi x/\lambda)$  and found that (1) there exists a critical wavelength  $\lambda_{cr}$  such that for smaller wavelengths the perturbation is stable and for larger ones it is unstable and (2) the value of the critical wavelength decreases appreciably with increase in the slipping velocity. As confirmed by simulations, proper resolution of high slip velocities during dynamic instabilities requires the statically estimated critical cell size  $h^*$  to be discretized by tens and sometimes (e.g., in the case of strong velocity weakening) even hundreds of cells  $h$ . Proper discretization is further discussed in sections 4 and 7 and by *Zheng and Rice* [1998].

The need for such fine discretization stems from the rate- and state-dependent friction laws that we discuss here. As already mentioned, these laws are supported by experimental evidence at low  $V$ , and their feature of state evolution over a slip distance  $L$  is supported by the concept of a characteristic slip required for renewal of the asperity contact population. The laws produce high slip velocities near the rupture tips and incorporate small characteristic slip distances to be resolved there, and hence require fine discretization in space and time. The discretization constraints may be possible to relax by using modified forms of the friction laws, for example, in which  $L$  in the law for  $d\theta/dt$  depends on  $V$  and increases significantly for the seismic range of  $V$ . Taking  $L$  proportional to  $V$  at high slip rates would be equivalent to having state evolve over a characteristic time (rather than over a characteristic slip distance), as for the velocity-weakening range of the ad hoc type of friction law used by *Shaw and Rice* [2000]. Such modifications and their influence on the qualitative features of the simulation results still

have to be explored.

#### 4. Variable Evolution Time Step

Simulating truly slow loading while capturing details of occasional rapid failures requires varying evolution time steps. Our time step selection criterion is based on two observations. First, we recognize that the slower the particle velocities in a rupture process are, the longer the time steps should become, and vice versa. Second, to assure proper integration of the constitutive law during the calculation, we would like the relative displacement in each time step to be small compared to the characteristic slip evolution distance  $L$ . To fulfill both of the above requirements, the time step from one updating of field variables (slip velocity, stress, etc.) to another, which we call the evolution time step  $\Delta t_{ev}$ , is chosen as

$$\Delta t_{ev} = \min[\xi_i L_i / V_i], \quad (11)$$

where  $L_i$ ,  $V_i$ , and  $\xi_i$  are the characteristic slip distance, the current slip velocity, and a prescribed parameter for the  $i$ th cell of the discretized domain (introduced earlier by  $[z_{i-1}, z_i]$ ,  $z_i = ih$ ,  $i = 1, 2, \dots, N_{ele}$ ,  $h = \lambda/N_{ele}$ ), respectively, and the minimum is sought over all the cells. The choice of parameters  $\xi_i$  depends on the constitutive law and stability considerations as explained below. Criterion (11) allows us to adjust the evolution time stepping during a simulation based on current slip velocities of the cells, so that slip in a time step does not exceed a fraction of the characteristic slip distance of the friction law, the fraction being prescribed by  $\xi_i$  for cell  $i$ . The adaptive time step  $\Delta t_{ev}$  can be enormously longer than the time for waves to propagate over the space domain during periods of slow, essentially quasi-static, loading, before unstable rupture begins, and can be very small during spontaneous failure, spanning up to 10 orders of magnitude in value in some of the simulations that we have done.

In between occurrences of dynamic ruptures, when slip rates are very small, we would like evolution time stepping to be as large as possible without compromising the algorithm accuracy and stability. A rather insightful constraint on the time steps at low slip rates can be derived by considering, as in our motivation for existence of the critical cell size, quasi-static stability of perturbed motion of a single cell with continuing steady sliding of the other cells at velocity  $V^*$ . If the grid is properly refined, then the perturbation on a single cell dies away, as we have already considered. Demanding that our time discretization preserves this property, we get a condition for the size of the time step allowed, which provides a constraint for  $\xi_i$ ,  $i = 1, 2, \dots, N_{ele}$  from (11). We derive this constraint in Appendix A by analyzing, as a simple



model case, explicit integration of the governing equations with a constant time step  $\Delta t$ . From that we obtain:

$$\Delta t < \frac{L}{V^*} \frac{A^*}{kL - (B^* - A^*)} \quad (12a)$$

if  $\chi > 0$  and

$$\Delta t < \frac{L}{V^*} \left( 1 - \frac{B^* - A^*}{kL} \right) \quad (12b)$$

if  $\chi < 0$ , where

$$\chi = \frac{1}{4} \left( \frac{kL}{A^*} - \frac{B^* - A^*}{A^*} \right)^2 - \frac{kL}{A^*} \quad (12c)$$

and  $A^*$  and  $B^*$  are given by

$$A^* = [V \partial \tau(V, \theta) / \partial V]_{V=V^*, \theta=\theta_{ss}(V^*)}, \quad (13)$$

$$A^* - B^* = [V d\tau_{ss}(V) / dV]_{V=V^*}.$$

As the derivation in Appendix A shows, constraints (12) are applicable to both steady-state velocity strengthening and steady-state velocity weakening with a sufficiently dense grid. For the latter case we have  $(B^* - A^*)/L = k_{cr}$  and  $k/k_{cr} = h^*/h > 1$ , and hence (12) can be rewritten in a more insightful form

$$\Delta t < \frac{L}{V^*} \frac{A^*}{(B^* - A^*)(h^*/h - 1)} \quad (14a)$$

if  $\chi > 0$  and

$$\Delta t < \frac{L}{V^*} \left( 1 - \frac{h}{h^*} \right) \quad (14b)$$

if  $\chi < 0$ , with

$$\chi = \frac{1}{4} \frac{B^* - A^*}{A^*} \left( \frac{h^*}{h} - 1 \right)^2 - \frac{h^*}{h}. \quad (14c)$$

Conditions (12) or (14) give an estimate of the required time stepping for low slip velocities; if these conditions are not met, cell-by-cell instabilities arise which either make the simulations impossible or corrupt the results.

Deriving these restrictions is an important development in the methodology, as they explained and eliminated many of the numerical difficulties that we had. Note that when the condition (12a) or (14a) is applicable (which is often the case since large  $h^*/h$  is required for proper space discretization), it implies that if we refine the grid (taking smaller cell size  $h$  and hence larger  $h^*/h$  in (14a) or larger  $k$  in (12a)), then we have to decrease the time stepping as well, even

in purely quasi-static phases of the analysis. The condition also reveals that if the direct effect  $A^*$  is decreased, then the time steps should be chosen smaller as well. That is why the efficiency of our algorithm, which relies on using long adaptive time steps during quasi-static loading periods, depends on the size of the positive direct effect (quantified by  $A^*$ ) for the quasi-static range of slipping velocities. Note that (12) or (14) are not applicable to the case  $A^* = 0$ , as their derivation (Appendix A) assumes nonzero  $A^*$ . Moreover, for a certain range of (very small) values of  $A^*$ , the inertial effects become comparable to the direct effect even for small sliding velocities and can no longer be ignored. The *Rice and Ruina* [1983] inertial analysis indicates that linearized perturbation to steady-state sliding of all wavelengths are unstable in the case of  $A^* = 0$ . In practice, it is possible to simulate a single dynamic event in the case with  $A^* = 0$ , apparently without numerical instability (A. Cochard, private communication, 1999), possibly due to the low rates of growth of the instability for the highest modes. However, the time stepping has to be so small that long deformation histories are excluded from consideration. Since our algorithm relies on using long adaptive time steps during quasi-static loading periods, it can be efficiently used only for the constitutive laws that exhibit the experimentally verified positive direct effect as in (8).

On the basis of conditions (12), we choose parameters  $\xi_i$ ,  $i = 1, 2, \dots, N_{ele}$  (used in selection of the evolution time steps (11)) as

$$\xi_i = \min \left\{ \frac{A_i^*}{kL_i - (B_i^* - A_i^*)}, \frac{1}{2} \right\} \quad (15a)$$

if  $\chi_i > 0$  and

$$\xi_i = \min \left\{ 1 - \frac{B_i^* - A_i^*}{kL_i}, \frac{1}{2} \right\} \quad (15b)$$

if  $\chi_i < 0$ , where

$$\chi_i = \frac{1}{4} \left( \frac{kL_i}{A_i^*} - \frac{B_i^* - A_i^*}{A_i^*} \right)^2 - \frac{kL_i}{A_i^*}. \quad (15c)$$

and subscript  $i$  denotes the value of the corresponding quantities for the cell  $i$  and  $k$  is the single-cell stiffness,  $\gamma\mu/h$ . The term  $1/2$  enters (15) to enforce the condition that for each cell the slip in every time step is not larger than half of the characteristic slip distance  $L_i$ .

The time step selection criterion (11) and (15) captures the essence of our variable time-stepping scheme, but in actual simulations we modify the criterion slightly to reconcile the variability in time steps with the necessity to compute convolution integrals, uniformly discretized in time. To store deformation histories in an efficient way, we introduce a time parameter,  $\Delta t_{min}$ , which is the minimum value of the

evolution time step allowed (and also the discretization interval for computing the convolution integrals, as explained in section 6). It is selected as a fraction of the time  $\Delta t_{\text{cell}}$  needed for elastic waves to traverse a spatial element, in the form

$$\Delta t_{\text{min}} = \beta_{\text{min}} \Delta t_{\text{cell}} = \beta_{\text{min}} h/c. \quad (16)$$

We insist that every time step we take be an integer multiple of  $\Delta t_{\text{min}}$  and not smaller than  $\Delta t_{\text{min}}$ . That is, we first compute the (tentative) time step  $\Delta t_{\text{ev}}$  using criterion (11) and (15) and then convert it into a multiple of  $\Delta t_{\text{min}}$ :

$$n_{\text{ev}} = \text{int} [\Delta t_{\text{ev}} / \Delta t_{\text{min}}], \quad (17)$$

$$\Delta \hat{t}_{\text{ev}} = \max \{ \Delta t_{\text{min}}, n_{\text{ev}} \Delta t_{\text{min}} \}.$$

The parameter  $\Delta t_{\text{min}}$  determines how fine our resolution in time is. To understand how to choose  $\Delta t_{\text{min}}$  (or  $\beta_{\text{min}}$  from (16)), let us consider how the evolution time step  $\Delta t_{\text{ev}}$  during dynamic instability relates to  $\Delta t_{\text{cell}} = h/c$ . We recognize from (1) that a characteristic slip velocity of order  $c\Delta\tau/\mu$  is induced by an abrupt dynamic stress drop  $\Delta\tau$ . Hence, to resolve the characteristic slip distance  $L$  of the friction law,  $\Delta t_{\text{ev}}$  has to be of the order  $L\mu/(c\Delta\tau)$ . At the same time, using our constraints (10) on the grid spacing  $h$ , we can express  $\Delta t_{\text{cell}} = h/c = h^*/(\eta c) = \gamma\mu L/(\eta c\Delta\tau_s)$ , where  $\Delta\tau_s = -\max[V d\tau_{\text{ss}}(V)/dV]$ . From the above formulae,  $\Delta t_{\text{ev}}/\Delta t_{\text{cell}}$  is comparable to  $\eta\Delta\tau_s/\gamma\Delta\tau$ . The constant  $\gamma$  is of order unity,  $\eta = h^*/h$  has to be tens or hundreds in order to properly discretize the critical cell size  $h^*$ , and the ratio  $\Delta\tau_s/\Delta\tau$  can be considerably smaller than unity. This suggests that the smallest required time step  $\Delta t_{\text{ev}}$  is comparable to  $\Delta t_{\text{cell}}$ .

We have found that if other parameters, most notably  $\eta = h^*/h$ , are chosen appropriately, the standard "sampling" choice  $\beta_{\text{min}}$  of 1/2, giving  $\Delta t_{\text{min}} = \Delta t_{\text{cell}}/2$ , produces stable and satisfactory results for the cases we considered;  $\beta_{\text{min}} = 1/4$  can also be successfully used in most cases. Note that (16) can be rewritten as

$$\Delta t_{\text{min}} = \beta_{\text{min}} h/c = \beta_{\text{min}} h^*/(\eta c). \quad (18)$$

If a better resolution in time is desired, it is often advantageous, for better stability and faster convergence of the results, to keep  $\beta_{\text{min}} = 1/2$  or  $1/4$  and to increase  $\eta = h^*/h$  (and hence  $N_{\text{ele}}$ ), rather than to keep  $\eta$  and to decrease  $\beta_{\text{min}}$ , even though increasing  $\eta$  is more costly in terms of computational time and memory. If there are numerical oscillations or other features in the simulation that point to an inadequate resolution, possibly in time, decreasing  $\beta_{\text{min}}$  without increasing  $\eta$  does not always solve the numerical problems

to the desired degree, whereas sufficient increase in  $\eta$  does, provided other parameters, such as elastodynamic time windows, replication period, etc., are chosen appropriately.

## 5. Updating Scheme: Advancing One Evolution Time Step

Let us consider how the values of field variables are updated over one evolution time step. We will use the velocity formulation without truncation in this section, for generality, and consider truncation of the convolution integrals in section 6. We continue the antiplane case with the domain  $\lambda$  discretized into cells  $[z_{i-1}, z_i]$ ,  $z_i = ih$ ,  $i = 1, 2, \dots, N_{\text{ele}}$ ,  $h = \lambda/N_{\text{ele}}$ . Suppose that the discretized values of slip  $\delta_i(t)$ , slip velocity  $V_i(t)$ , state variable  $\theta_i(t)$ , stress  $\tau_i(t)$ , and state rate  $\dot{\theta}_i(t)$  at cell centers are known at time  $t$  for all  $i = 1, 2, \dots, N_{\text{ele}}$  and that the slip velocity history is known for all prior time  $t'$ ,  $0 < t' < t$ , where  $t = 0$  is the beginning of the deformation process considered. When using the spectral formulation, we also assume that the Fourier coefficients  $D_n(t)$  of the slip distribution are known at time  $t$  and note that the velocity history need only be available in the Fourier domain, as the values of  $\dot{D}_n(t')$  for  $0 < t' < t$ . To advance the field values by one evolution time step and to determine all the quantities just mentioned at the end of that step, we proceed in the spirit of a second-order Runge-Kutta procedure as follows:

1. Determine the evolution time step  $\Delta t = \Delta \hat{t}_{\text{ev}}$  to be made using criterion (11) and (15) - (17).
2. Make first predictions of the values of slip and state variable at time  $t + \Delta t$ , based on known values at  $t$ , as

$$\begin{aligned} \delta_i^*(t + \Delta t) &= \delta_i(t) + \Delta t V_i(t), \\ \theta_i^*(t + \Delta t) &= \theta_i(t) + \Delta t \dot{\theta}_i(t). \end{aligned} \quad (19)$$

3. Make a corresponding first prediction  $f_i^*(t + \Delta t)$  of the functional, using slip prediction (19) and treating the slip rates as if they were constant through the time step  $\Delta t$  and equal to  $V_i(t)$ . To implement this in the spectral formulation, we first compute the Fourier coefficients of  $V_i(t)$  and  $\delta_i^*(t + \Delta t)$ . To represent FFT operations, we shift the coordinate origin so that the cell centers are at  $z_j = jh$ ,  $j = 1, 2, \dots, N_{\text{ele}}$ , and define  $R_{nj} = e^{-2\pi i n j / N_{\text{ele}}} / N_{\text{ele}}$  and  $R_{jn}^{-1} = e^{2\pi i n j / N_{\text{ele}}}$ , where  $i = \sqrt{-1}$  (unless used as sub-

script), to get

$$\dot{D}_n(t) = \sum_{i=1}^{N_{ele}} R_{ni} V_i(t), \quad (20a)$$

$$D_n^*(t + \Delta t) = D_n(t) + \Delta t \dot{D}_n(t).$$

Then, using (4), we get the Fourier coefficients of the prediction of the functional

$$F_n^*(t + \Delta t) = -\frac{\mu |k_n|}{2} \left[ D_n^*(t + \Delta t) - \int_{\Delta t}^{t+\Delta t} W(|k_n|ct') \dot{D}_n(t + \Delta t - t') dt' - \dot{D}_n(t) \int_0^{\Delta t} W(|k_n|ct') dt' \right]. \quad (20b)$$

Within the brackets of (20b), the second term can be computed since the slip velocity history is known. The third term is an approximation of the convolution on the time interval corresponding to the current time step. We then obtain the prediction of the functional through an inverse FFT as

$$f_i^*(t + \Delta t) = \sum_{n=-N_{ele}/2}^{N_{ele}/2} R_{in}^{-1} F_n^*(t + \Delta t). \quad (20c)$$

4. Find predicted slip rates  $V_i^*(t + \Delta t)$  corresponding to the predicted state  $\theta_i^*(t + \Delta t)$  from (19) and functional  $f_i^*(t + \Delta t)$  from the last stage. This is done by equating stress (1) to the strength (7a) to get

$$\begin{aligned} \tau_i^o(t + \Delta t) + f_i^*(t + \Delta t) - \frac{\mu}{2c} V_i^*(t + \Delta t) \\ = \psi(V_i^*(t + \Delta t), \theta_i^*(t + \Delta t)) \end{aligned} \quad (21)$$

and then solving (21) for  $V_i^*(t + \Delta t)$ . We find  $V_i^*(t + \Delta t)$  using Newton-Rhapson search with  $V_i^*(t)$  as the first guess. Once  $V_i^*(t + \Delta t)$  are obtained, the corresponding state rates can be readily found from (7b) as  $\dot{\theta}_i^*(t + \Delta t) = \varphi(V_i^*(t + \Delta t), \theta_i^*(t + \Delta t))$ .

5. Calculate the final prediction of slip and state variable at time  $t + \Delta t$  by

$$\delta_i^{**}(t + \Delta t) = \delta_i(t) + \frac{\Delta t}{2} [V_i(t) + V_i^*(t + \Delta t)], \quad (22)$$

$$\theta_i^{**}(t + \Delta t) = \theta_i(t) + \frac{\Delta t}{2} [\dot{\theta}_i(t) + \dot{\theta}_i^*(t + \Delta t)].$$

(The superscript double asterisks could be dispensed with at this point, but is useful for comparison to other methods.)

6. Make a corresponding prediction  $f_i^{**}(t + \Delta t)$  of the functional, using the  $\delta_i^{**}(t + \Delta t)$  and treating the slip rates as if they were constant throughout the time step at  $(V_i(t) + V_i^*(t + \Delta t))/2$ , consistently with updating slip in stage 5. The specific steps are analogous to (20) in stage 3. Note that the second term in the brackets of (20b) will be the same in this stage, and hence it can be computed just once, in stage 3, and stored for use here. This is computationally very advantageous, since this term incorporates most of the convolution evaluation.

7. Make final predictions  $V_i^{**}(t + \Delta t)$  and  $\dot{\theta}_i^{**}(t + \Delta t)$  of the slip rate and state rate, similar to stage 4, using the new predictions  $\theta_i^{**}(t + \Delta t)$  for the state variable from (22) and the result  $f_i^{**}(t + \Delta t)$  of the last stage for the functional. That is, solve the equation like (21), but with superscripts double asterisks, to find  $V_i^{**}(t + \Delta t)$ , and then compute  $\dot{\theta}_i^{**}(t + \Delta t)$  from (7b).

8. Declare the values of field quantities  $\delta_i(t + \Delta t)$ ,  $\theta_i(t + \Delta t)$ ,  $V_i(t + \Delta t)$ ,  $\dot{\theta}_i(t + \Delta t)$  to be equal to the predictions with the superscript double asterisks. Compute the corresponding values of stress  $\tau_i(t + \Delta t)$ , if needed, from (7a). Store slip velocity history for the time interval  $[t, t + \Delta t]$ , to be used in future convolution evaluations, as  $V_i(t') = [V_i(t) + V_i^*(t + \Delta t)]/2$  for  $t'$  in  $[t, t + \Delta t]$ . In the spectral formulation, store instead the history of the Fourier coefficients as  $\dot{D}_n(t') = (\dot{D}_n(t) + \dot{D}_n(t + \Delta t))/2$  for  $t'$  in  $[t, t + \Delta t]$ , since they are actually used in the convolutions, and set  $D_n(t + \Delta t) = D_n^{**}(t + \Delta t)$ . Finally, return to stage 1 to advance through the next time step.

This scheme is second-order accurate in  $\Delta t$  for the slip and state variable, assuming that the predictions of the functional are computed accurately enough (N. Lapusta, Ph.D. thesis in preparation, 2000). We use the midpoint integration scheme to compute the convolution integrals. Note that while it is important to know the order of accuracy of an updating scheme, the actual performance also depends on other things, such as stability characteristics, balance of terms that achieves most error cancellation, etc. Ultimately, the most important thing is the ability of a scheme to produce numerically stable simulations with results convergent through space grid reduction and better time resolution, which can often be checked only by actually doing the simulation.

Earlier studies used different updating schemes that provided a foundation for the development of the scheme described here. The scheme used by *Rice and Ben-Zion* [1996] and *Ben-Zion and Rice* [1997] proceeds through stages 1 to 4 and declares the values of the field variables at time  $t + \Delta t$  to be equal to the predictions  $\delta_i^*(t + \Delta t)$ ,  $\theta_i^*(t + \Delta t)$ ,  $V_i^*(t + \Delta t)$ , and  $\dot{\theta}_i^*(t + \Delta t)$ , all of which are first-order ac-

curate (N. Lapusta, Ph.D. thesis in preparation, 2000). The only departure in this "incomplete" scheme is in stage 2, where the value of the state was computed using not (19), but through exact integration of (7b) assuming that the slip velocity was constant in time throughout the step. Another updating scheme was originally developed by *Morrissey and Geubelle* [1997] for the case of constant evolution time steps and constitutive laws without state variables and described by them as a "semi-implicit velocity formulation", "with delay", "discretized kernel", and "convolutions by trapezoid rule". It incorporates steps similar to stages 1-5, and then uses predictions  $\delta_i^{**}(t + \Delta t)$ ,  $\theta_i^{**}(t + \Delta t)$ ,  $V_i^*(t + \Delta t)$ , and  $\theta_i^*(t + \Delta t)$  as the values of the field variables at time  $(t + \Delta t)$ . It also approaches differently the convolution evaluation, using a trapezoidal rule and delay in kernel. The delay in kernel, discussed in detail by *Morrissey and Geubelle* [1997], was introduced as an empirical step by *Cochard and Madariaga* [1994] to smooth numerical oscillations in slip velocity right behind the rupture front, but at the cost of reducing the slip velocities at the tips of the propagating disturbances.

The present scheme performs better than both of the above mentioned schemes in the cases that we considered. It does not use the delay, captures more accurately the (high) slip velocities at the rupture tips, and has essentially the same stability performance. However, it is more costly in terms of the computational time (but not memory, which is often the primary limitation), mostly because it uses another pair of FFT transforms at the stage 6. Each of the FFT transforms requires  $O(N_{\text{ele}} \log_2(N_{\text{ele}}))$  floating point operations. The other time-consuming computation is the evaluation of convolution integrals, which, as considered in section 6, requires from  $O(N_{\text{ele}} \log_2(N_{\text{ele}}))$  to  $O(N_{\text{ele}}^2)$  operations, depending on the truncation procedure used. Clearly, if the truncation scheme used requires  $O(N_{\text{ele}}^2)$  operations, then the extra FFTs do not make much difference in terms of the cpu time. However, if the more efficient truncation procedure can be used in the problem at hand, then the number of operations for the FFTs and for the convolutions can have comparable orders of magnitude, in which case the advantage of the full scheme 1-8 has to be weighted against the increase in the computational time.

## 6. Evaluation of the Truncated Convolution Integrals

Let us consider the evaluation of the truncated convolution integrals in the velocity formulation (1), (2), and (6). The elastodynamic window  $T_w$  (introduced in section 2) can be selected the same for all Fourier modes, or it can be mode-dependent. We have studied both approaches. Let

$T_w(n) = T_w(|n|)$  denote the length of the elastodynamic time window for mode  $n$ . Keeping the window the same for all Fourier modes simplifies the procedure, but it is not a very efficient choice, for the following reason. The argument of the convolution kernel,  $|k_n|c\theta = 2\pi c|n|\theta/\lambda$ , depends on the mode number  $n$  and varies in the ranges  $[0, 2\pi cT_w(1)/\lambda]$  for the lowest (spatially nonuniform) mode  $|n| = 1$  and  $[0, \pi N_{\text{ele}}cT_w(N_{\text{ele}}/2)/\lambda]$  for the highest mode  $|n| = N_{\text{ele}}/2$ . We call the length of these ranges kernel windows  $K_w(n)$ , so that

$$K_w(n) = |n|(2\pi c/\lambda)T_w(n). \quad (23)$$

If  $T_w(n)$  is the same for all  $n$ , then  $K_w$  for the highest mode is  $(N_{\text{ele}}/2)$  times longer than for the lowest mode. Since the convolution kernel decays (Figure 1) and  $N_{\text{ele}}$  is usually a large number (spanning values from 512 to 65536 in the simulations we have done), much of the computation for the highest modes has negligible contribution.

To save computational time and memory, we examined use of time windows  $T_w$  which are mode-dependent and significantly shorter for the higher modes. In the current implementation, two parameters determine the window sizes. One of them is  $T_w(1)$ , the time window for the lowest mode  $|n| = 1$ , and the other is  $q_w$ , the ratio of the  $K_w$  for the highest and the lowest modes. Once these parameters are selected, the  $K_w$  for all the modes are determined as

$$K_w(1) = (2\pi c/\lambda)T_w(1), \quad K_w(N_{\text{ele}}/2) = q_w K_w(1), \quad (24)$$

$$\begin{aligned} K_w(n) &= K_w(1) + \frac{K_w(N_{\text{ele}}/2) - K_w(1)}{N_{\text{ele}}/2 - 1} (|n| - 1) \\ &= (2\pi c/\lambda)T_w(1) \left( 1 + \frac{q_w - 1}{N_{\text{ele}}/2 - 1} (|n| - 1) \right). \end{aligned}$$

That is, the kernel window  $K_w(N_{\text{ele}}/2)$  for the highest mode is  $q_w$  times longer than the window  $K_w(1)$  for the lowest mode, and the kernel windows for the modes in between vary linearly with  $|n|$ . The corresponding time windows  $T_w$  can be found from (23). Note that for  $q_w = N_{\text{ele}}/2$ , this approach is equivalent to the one with the constant  $T_w$ . The advantage arises from the fact that far smaller  $q_w$  can be used; acceptable values can be as low as 4 for some problems, including our implementation examples in section 7.

We determine the parameters  $T_w(1)$  and  $q_w$  by trial and error, starting with an educated guess and then comparing the results with the ones for smaller and larger values of both parameters, until convergence is reached. A useful parameter for making the initial  $T_w(1)$  guess is the time  $T_\lambda$  for elastic waves to propagate through the domain of size  $\lambda$  treated

in the spectral formulation. In general,  $T_w(1) = T_\lambda$  is a good initial guess. We select  $T_w(1)$  in the form

$$T_w(1) = \eta_w T_\lambda = \eta_w \lambda / c, \quad (25)$$

where, from our experience,  $\eta_w$  of 1 to 4 are sufficient values for most problems.

In order to use standard procedures for computing the convolution integrals, we discretize  $T_w(n)$  using the (constant) time interval  $\Delta t_{\min}$  introduced in (16). Computing the field values using evolution time steps  $\Delta \hat{t}_{\text{ev}}$  from (17), which are multiples of  $\Delta t_{\min}$ , we store the field values needed for convolution evaluation on a uniform time grid of spacing  $\Delta t_{\min}$ , repeating each value  $\Delta \hat{t}_{\text{ev}} / \Delta t_{\min}$  times. The lengths of elastodynamic time windows  $T_w(n)$  determine how many values have to be stored from the current value back in time for each Fourier mode. This array of the stored deformation history contains the values needed for the discretized convolution in the useful format (spaced by  $\Delta t_{\min}$ ), which simplifies and speeds up the computation. At each evolution time step, the array of stored history is updated by writing newly computed values over the values that have moved outside the elastodynamic time windows  $T_w(n)$ . Hence only a fraction of the array is typically updated at each time step, and the assignment for each value in the array is done only once. In earlier implementations [Zheng *et al.*, 1995; Rice and Ben-Zion, 1996; Ben-Zion and Rice, 1997], the field values were stored at (variable) evolution time steps, and then time-consuming search and mapping routines were employed in every time step to transfer values from the set of the evolution time steps to the uniformly spaced array required for the calculation of the convolution integrals. Obtaining this array was the most expensive part of the computation, being up to 10 times more time-consuming than the multiplications needed to evaluate convolutions. The current procedure reduces the cpu time for updating the array of the stored history by orders of magnitude, so that the cpu time for computing convolution integrals is essentially determined by the cpu time required to perform multiplications.

Let us estimate the order of magnitude of this latter cpu time. When the time windows  $T_w$  are the same for each Fourier mode, they are taken to be of the order of the time  $T_\lambda$  for elastic waves to propagate through the domain of interest. In this case, the convolution evaluation, at each time step, requires  $O(N_{\text{ele}})$  floating point operations for each Fourier mode and  $O(N_{\text{ele}}^2)$  operations altogether (for all modes). In the case of  $T_w$  dependent on the Fourier modes according to (23)-(24), an upper bound on the number of operations for the mode  $n$  scales as  $O(q_w N_{\text{ele}} / |n|)$ , and the number of operations for all modes is at most  $O(q_w N_{\text{ele}} \ln(N_{\text{ele}}))$ . Since the typical values for  $N_{\text{ele}}$  are thousands to tens of thousands, and  $q_w$  can be as small as 4, the reduction in the

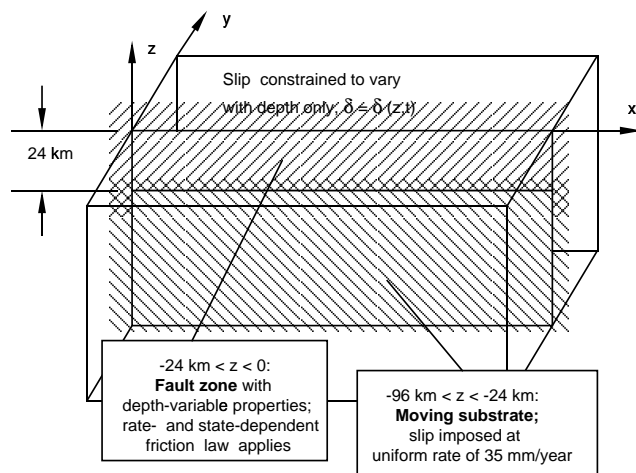
overall cpu time due to the mode-dependent  $T_w$  is very significant. The analogous reduction arises in memory requirements, as much less deformation history has to be stored for higher modes.

## 7. Implementation Example

### 7.1. Formulation of 2-D Model and its Response

To demonstrate how the ideas outlined in the previous sections are combined to produce long-duration simulations, let us consider elastodynamic response of a 2-D depth-variable fault model (Figure 2), to which earlier implementations of related procedures have already been applied [Rice and Ben-Zion, 1996; Ben-Zion and Rice, 1997]. In this model [Rice, 1993], a vertical strike-slip fault with depth-variable properties is embedded in an elastic half-space. The fault is driven below depth  $Z_{\text{depth}} = 24$  km with a plate rate of  $V_{\text{pl}} = 35$  mm/yr. In the shallower zone, governed by a constitutive law, the slip  $\delta(z, t)$  is calculated as a function of depth  $z$  and time  $t$  (variations with along-strike distance  $x$  are not included in this 2-D model).

This model can be mathematically described as the antiplane problem discussed in previous sections. It proves convenient to express the formulae in terms of variables  $(\delta(z, t) - V_{\text{pl}}t)$  and  $(V - V_{\text{pl}})$ , in which case  $\tau^o(z, t)$ , the stress which would act if the plane  $y = 0$  were constrained against any slip, becomes independent of time and equal to the initial stress  $\tau^o(z)$ . Hence, following relations (1) and



**Figure 2.** A vertical strike-slip fault in an elastic half-space (like in the work by Rice [1993]).

(2), we write

$$\tau(z, t) = \tau^o(z) + f(z, t) - \frac{\mu}{2c} (V(z, t) - V_{\text{pl}}), \quad (26)$$

$$\delta(z, t) - V_{\text{pl}}t = \sum_{n=-N_{\text{ele}}/2}^{N_{\text{ele}}/2} D_n(t) e^{ik_n z/\lambda}, \quad (27)$$

$$f(z, t) = \sum_{n=-N_{\text{ele}}/2}^{N_{\text{ele}}/2} F_n(t) e^{ik_n z/\lambda},$$

with the relation between  $D_n(t)$  and  $F_n(t)$  given by (6). We use  $\mu = 30$  GPa and  $c = 3$  km/s.

We select the replication distance  $\lambda$  in the  $z$  direction in the following way. In addition to the region  $[-Z_{\text{depth}}, 0]$  where we wish to simulate slip, we include in  $\lambda$  the substrate (mantle) region  $[-4Z_{\text{depth}}, -Z_{\text{depth}}]$ , where a plate velocity of  $V = V_{\text{pl}} = 35$  mm/yr is imposed. That means  $\delta(z, t) - V_{\text{pl}}t = 0$  there, so FFT sample points in that part of the domain create zero padding. Finally, to model the free surface at  $z = 0$ , we map the slip and slip velocity from the region  $[-4Z_{\text{depth}}, 0]$  to the region  $[0, 4Z_{\text{depth}}]$  as even functions. Hence, the spatial domain to consider is  $[-4Z_{\text{depth}}, 4Z_{\text{depth}}]$  in the  $z$  direction, and  $\lambda = 8Z_{\text{depth}}$  or 192 km for  $Z_{\text{depth}} = 24$  km. As noted before, because of the Fourier series representation (27), we actually solve a problem where this domain is periodically repeated along  $z$ , but the zones of potential rapid slip accumulation, which coincide with the steady-state velocity-weakening regions of  $a - b < 0$ , are separated enough (by at least  $6Z_{\text{depth}}$ ) to prevent significant influence of spatial replications on each other.

For our examples here, we take the constitutive law in one of the standard laboratory-derived forms of rate- and state-dependent friction (7) with the Dieterich-Ruina version of state variable evolution

$$\tau(z, t) = \bar{\sigma}(z) \left[ f_o + a(z) \ln \frac{V(z, t)}{V_o} + b(z) \ln \frac{V_o \theta(z, t)}{L(z)} \right], \quad (28a)$$

$$\frac{\partial \theta(z, t)}{\partial t} = 1 - \frac{V(z, t) \theta(z, t)}{L(z)}, \quad (28b)$$

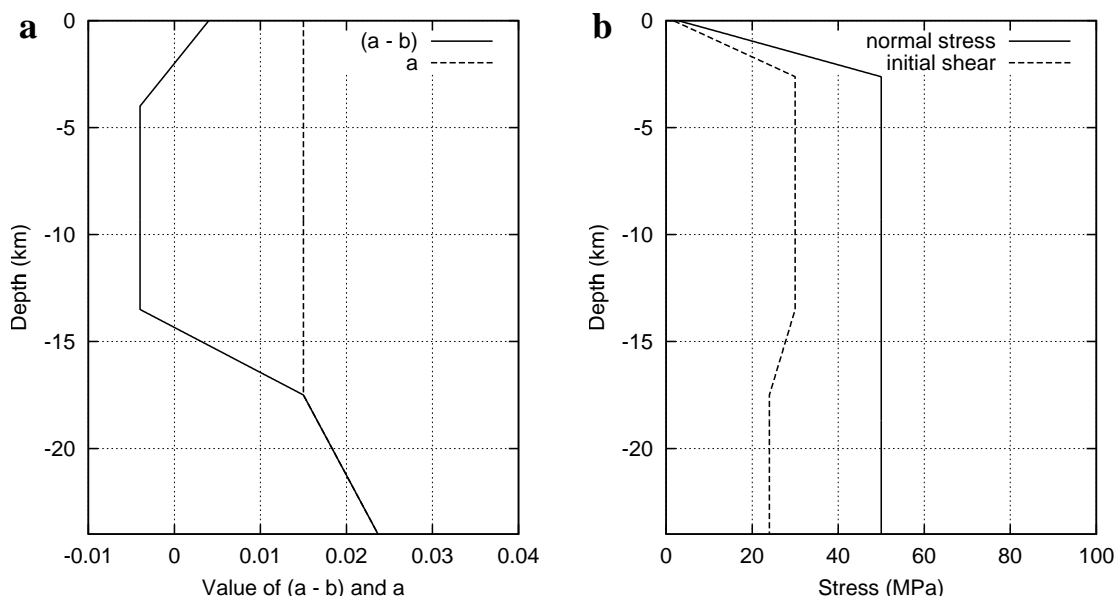
except that we regularize the first of these near  $V = 0$  as discussed below and, to allow  $V$  of either sign, replace  $V(z, t)$  with  $|V(z, t)|$  in (28b). In (28),  $\bar{\sigma}$  is the effective normal stress,  $f_o = 0.6$  is the value of friction coefficient at the reference velocity  $V_o = 1 \mu\text{m/s}$ ,  $a$  and  $b$  are frictional parameters, and  $L$ , as before, is the characteristic slip distance. As (28) indicates,  $\bar{\sigma}$ ,  $a$ ,  $b$ , and  $L$  vary with depth but not with time. Examples of dynamic modeling with addition of

power law creep at depth and other features are given by *Rice and Ben-Zion* [1996] and *Ben-Zion and Rice* [1997]. The assumed variation of  $a$  and  $b$  with depth, like in the work by *Rice* [1993], is shown in Figure 3a. This is consistent with the experimentally determined temperature dependence of  $(a - b)$  by *Blanpied et al.* [1991, 1995] for granite under hydrothermal conditions, as mapped by them into a depth variation based on a San Andreas fault geotherm. Variation of  $L$  with depth is discussed below. The effective normal stress  $\bar{\sigma}$  is assumed in this example to vary with depth in a way that incorporates high fluid over pressurization at depth, according to  $\bar{\sigma} = \min[2.8 + (18z/\text{km}), 50]$  MPa. In this distribution (shown in Figure 3b),  $\bar{\sigma}$  is equal to overburden minus hydrostatic pore pressure at shallow depth (up to about 2.6 km), with transition to lithostatic pore pressure gradient with 50 MPa offset at depth. Figure 3b also shows the initial stress, which is the same for all the cases considered here.

Since the friction law (28) is a particular (and widely used) version of (7), all our conclusions from sections 3 and 4 hold with  $A^* = a\bar{\sigma}$ ,  $B^* = b\bar{\sigma}$ . The critical stiffness (9a) becomes  $k_{\text{cr}} = \bar{\sigma}(b - a)/L$ . To find the critical cell size  $h^*$  corresponding to (10), we need to determine the coefficient  $\gamma$  for our present model in the expression for the single-cell stiffness  $k = \gamma\mu/h$ . As mentioned in section 3,  $\gamma = \pi/4 \approx 0.79$ . That  $\gamma$  was determined by performing the static elastic spectral analysis of unit slip at a single FFT sample point (and all its replications) and equating the resulting stress reduction there to  $\gamma\mu/h$ . When we ignore the presence of a free surface, the calculation is straightforward and gives  $\gamma = \pi/4 + 2/N_{\text{ele}}$  (where  $2/N_{\text{ele}}$ , which arises from the periodic replication, can be ignored for the large values of  $N_{\text{ele}}$  we consider). A full calculation, which we did numerically, also includes unit slip in the mirror cell and in all its periodic replicates. It resulted in values of  $\gamma$  very close to  $\pi/4$  for all the cells except for the one adjacent to the free surface, where  $\gamma = 0.47$ . Using  $\gamma = \pi/4$ , we estimate the critical cell size by

$$h^* = \frac{\pi}{4} \frac{\mu L}{\bar{\sigma}(b - a)}. \quad (29)$$

A drawback of the logarithmic form (28a) is that the stress is not defined for  $V = 0$ . The logarithmic form was derived from purely empirical considerations to match experimental observations [*Dieterich*, 1979, 1981; *Ruina*, 1983]. However, it has a theoretical basis, in that such a form would result if the direct velocity effect is due to stress biasing of the activation energy in an Arrhenius rate process at contact junctions, at least in the range for which forward microscopic jumps, in the direction of shear stress, are overwhelmingly more frequent than backward jumps. Such interpretation seems implicit in the work by *Chester*



**Figure 3.** (a) Depth-variable distribution of frictional parameters  $(a - b)$  and  $a$  (like in the work by Rice [1993]), consistent with the measured temperature and inferred depth variation of  $(a - b)$  of Blanpied *et al.* [1991, 1995] for granite under hydrothermal conditions. (b) Depth-variable distribution of the effective normal stress  $\bar{\sigma}$  (solid line) and initial shear stress  $\tau^o(z)$  (dashed line).

and Higgs [1992] and Chester [1994] and is more explicitly proposed by Brechet and Estrin [1994] and Baumberger [1997]. To account in a simple way for backward jumps, which could not be neglected near  $V = 0$ , we solve (28a) for  $V$ , identifying the factor  $\exp(\tau/a\bar{\sigma})$ , which then appears as the stress biasing of forward jumps, and replace it by  $[\exp(\tau/a\bar{\sigma}) - \exp(-\tau/a\bar{\sigma})]$  to account for backward jumps too. This procedure, used by Rice and Ben-Zion [1996] and Ben-Zion and Rice [1997], replaces (28a) with the regularized form

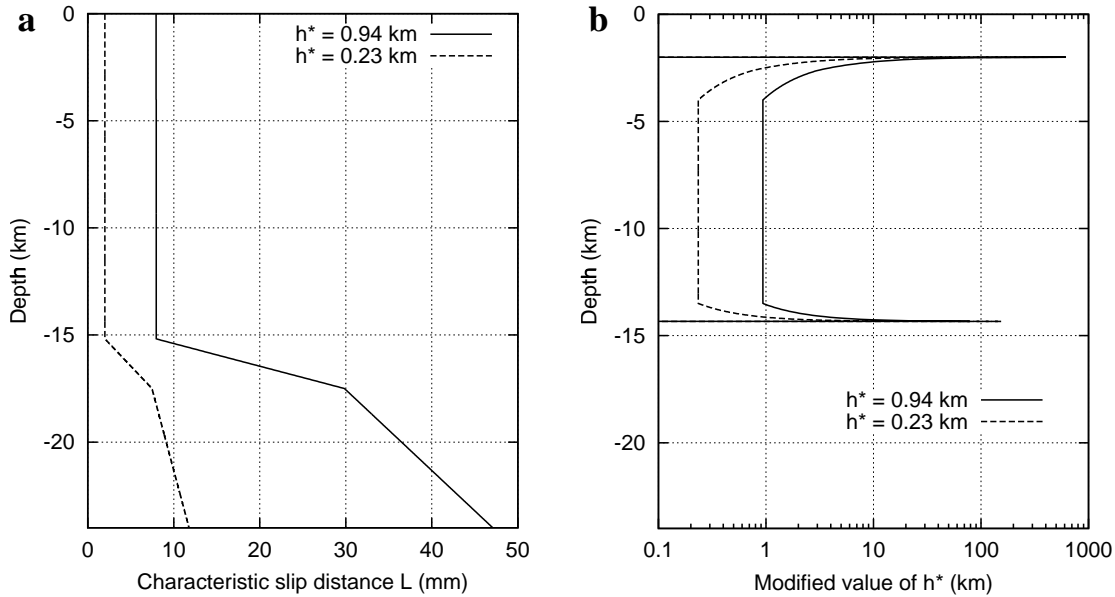
$$\tau(z, t) = a\bar{\sigma}(z, t) \times \operatorname{arcsinh} \left[ \frac{V}{2V_o} \exp \left( \frac{f_o + b \ln(V_o\theta/L)}{a} \right) \right]. \quad (30)$$

During forward sliding at rates of order  $V_o$  the modification to  $\exp(\tau/a\bar{\sigma})$  is of order  $\exp(-2f_o/a)$  or less, where  $f_o/a \approx 40$ , and so this is a negligible change from (28a).

As discussed before, the critical cell size  $h^*$  from (29) is a very important parameter, both for the physics and numerics of the problem. In principle,  $h^*$  can vary with depth, since it is determined by the depth-variable frictional properties and stress. In all our examples, however, we attempt to make  $h^*$  uniform throughout the velocity-weakening depth range. The motivation is that we have a uniform computa-

tional grid and wish to keep  $h^*/h$  uniformly high for good numerical resolution. As is well known, it is not presently feasible to do computations with values of  $h^*$  chosen in consistency with laboratory values of  $L$  (which are typically in the few micron range and give  $h^*$  in the range of 1 m). We thus take larger  $L$  and  $h^*$  but do keep  $h^*$  small compared to other feature sizes in the model, such as the seismogenic depth, which in our model corresponds to the steady-state velocity weakening region, extending over  $\approx 12$  km. We prescribe  $h^*$  in the simulations, using values approximately equal to 0.94 km and  $0.94/4 = 0.235$  km in the examples shown.

The distribution of the characteristic slip distance  $L$  is assigned from (29) based on the desired  $h^*$  for cells with steady-state velocity weakening. For the above values of  $h^*$ , this results in  $L$  of order of millimeters to tens of millimeters, much larger than laboratory values but, as explained, needed to keep the size of  $h^*$  possible to resolve. Such selection of  $L$  has to be adjusted in the regions close to transition from the velocity-weakening to the velocity-strengthening friction, where, owing to near-zero values of  $(b - a)$ , it results in very small values of  $L$ . Resolution of these values would require extreme refinement of time stepping during dynamic rupture, when slip velocities are large, as follows from the time selection criterion (11). Thus we increase  $L$  in those regions, effectively changing (increasing) the nucle-



**Figure 4.** (a) Depth-variable distribution of the characteristic slip distance  $L$  of the friction law, for the two cases considered. (b) Modification in the critical cell size  $h^*$  caused by the adjustments (increases) to  $L$  in the regions next to the transition from the velocity-weakening to the velocity-strengthening friction;  $h^*$  is defined only for the velocity-weakening part of the fault zone.

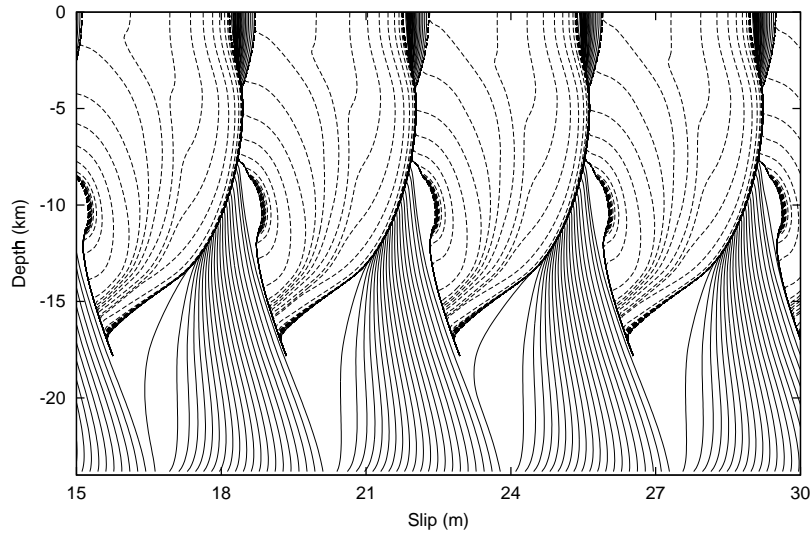
ation size there. As for the velocity-strengthening regions, keeping a particular value of  $h^*$  is not of a concern, since nucleation cannot happen there. However, it is the resolution of  $L$  in the velocity-strengthening regions that usually controls the size of time steps during essentially quasi-static phases of deformation. This is not surprising since during the quasi-static phases the velocity-weakening regions are stuck with near-zero velocities, while the velocity-strengthening regions are creeping with (much larger) slip velocities close to the plate velocity. Hence, while assigning  $L$  in the velocity-strengthening regions, it is practical to take into consideration time step constraints (12). Keeping in mind the restrictions discussed and aiming for a simple and continuous distribution of  $L$ , we use, for the examples in this paper, the distribution of  $L(z)$  shown in Figure 4a. Such an assignment of  $L$  keeps  $h^*$  constant and unmodified from 13.5 to 4 km depth. Figure 4b demonstrates the modification of  $h^*$  caused by the adjustments in  $L$ .

Before we consider choices of numerical parameters, let us have a look at the response this model produces and our simulations are able to capture. Figures 5 and 6 show parts of the slip accumulation  $\delta$  for the cases  $h^* = 0.94$  and 0.235 km, respectively. Figure 7 shows the maximum slip velocity histories for these two cases. From the data, as well as from the assorted slip velocity output, we notice that the velocity-strengthening region at the bottom of the fault is creep-

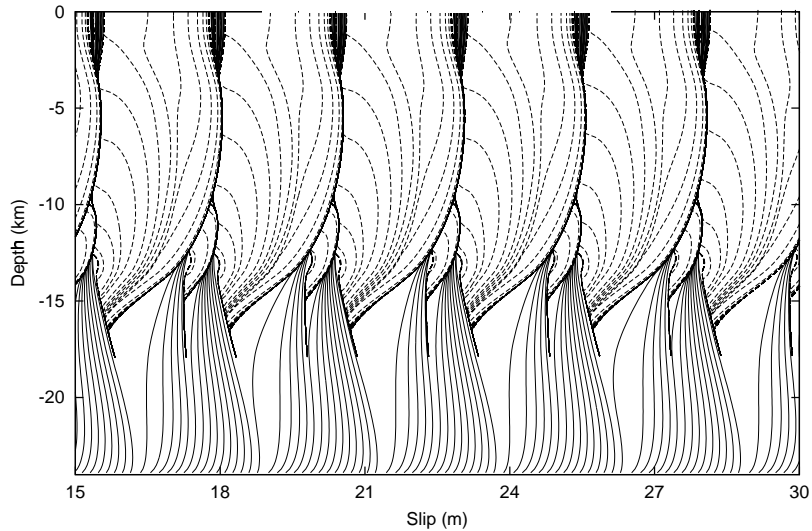
ing, with roughly the plate velocity of 35 mm/yr, whereas the velocity-weakening region accumulates slip through dynamic failure events. The velocity-strengthening region near the free surface is thin and gets broken by strong ruptures coming from the bottom of the fault segment, but it also exhibits some creeping. Even on these plots, nucleation zones can be distinguished (especially for the case  $h^* = 0.94$  km), corresponding to several dashed lines plotted on top of each other. The slip accumulates quasi-statically at those regions for a long time, then the slip accelerates and dynamically expands from there. Notice that the nucleation size scales with  $h^*$ , and in both cases the nucleation size is  $\approx 3.5$  times larger than  $h^*$ . The model earthquakes generally nucleate at the bottom transition between the velocity-weakening and velocity-strengthening regions, because it is there that the creeping region transfers stresses to the seismogenic depth, loading it up. By looking at the distance between the tips of the dashed lines, which are 1 s apart in time, we can estimate the rupture propagation velocities. These range from slower ones right after the nucleation up to almost the shear wave speed of 3 km/s further in the rupture development. Many of the model earthquakes are large and reach the free surface, sending a wave of slip back to depth.

The provided examples show that our simulation algorithm deals very well with slow loading of the fault with the equivalent of the millimeters per year plate rate, slow

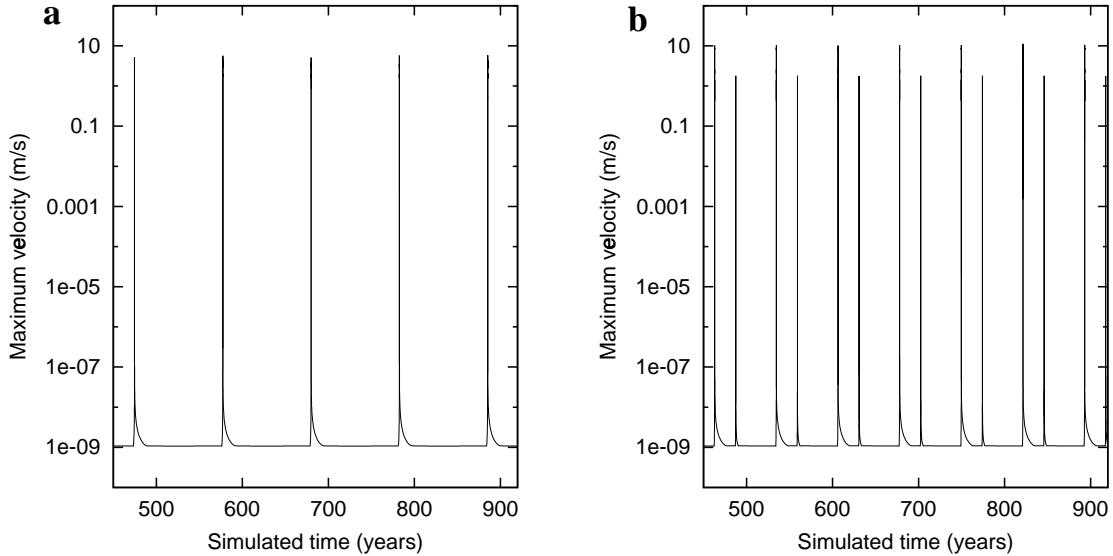




**Figure 5.** Accumulation of slip versus depth for the case  $h^* = 0.94$  km,  $h^*/h = 40$ . The solid lines are plotted every 5 years. The dashed lines are plotted above 18 km depth every second if the maximum velocity anywhere on the fault exceeds 0.001 m/s. The model response consists of large, essentially periodic events rupturing the whole fault. The nucleation size is  $\approx 3.2$  km.



**Figure 6.** Accumulation of slip versus depth for the case  $h^* = 0.235$  km,  $h^*/h = 40$ . The solid lines are plotted every 5 years. The dashed lines are plotted above 18 km depth every second if the maximum velocity anywhere on the fault exceeds 0.001 m/s. The model response consists of a repeated pair of larger and smaller events. The nucleation size is  $\approx 0.8$  km. In comparison with Figure 5, note the difference in the system behavior and the change in the nucleation size as  $h^*$  is changed.



**Figure 7.** Maximum slip velocity on the fault as a function of time for (a)  $h^* = 0.94$  km and (b)  $h^* = 0.235$  km. Individual rupture events collapse onto a straight line on this timescale of hundreds of years. The maximum slip velocities reached during dynamic ruptures are of the order of 10 m/s for larger events and 1 m/s for smaller events in the Figure 7b. The maximum slip velocity in between the dynamic events is  $\approx 10^{-9}$  m/s, which corresponds to the plate velocity of 35 mm/yr. Read 1e-05 as  $10^{-5}$ .

event nucleation, and dynamic slip with kilometers per second rupture velocities and meters per second slip velocities. The results shown are well-resolved numerically. We now discuss the choices of numerical parameters that produce these simulations.

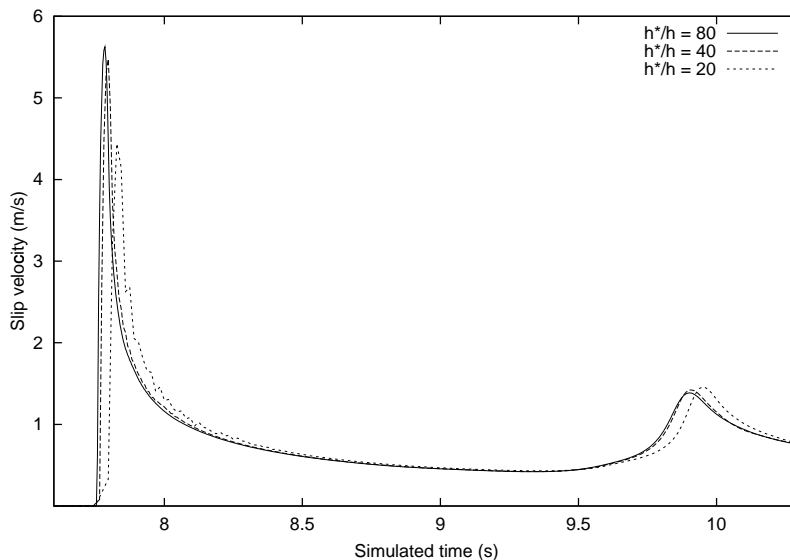
## 7.2. Parameter Selection for Well-Resolved Simulations

As explained in sections 3 and 4, the choice of the parameter  $\eta = h^*/h$  is crucial for the simulation stability, accuracy, and tractability. Together with the selection of  $h^*$  and  $\lambda$ , this parameter determines the number  $N_{\text{ele}}$  of the discretization points required (and hence the problem size) and the cell size  $h$  through  $N_{\text{ele}} = \lambda/h = \eta\lambda/h^*$  and  $h = h^*/\eta$ . If  $\beta_{\text{min}}$  from the expression for the minimum time step (16) is kept unchanged, the parameter  $\eta$  can also be considered as controlling the smallest time discretization allowed. For the examples here, we find that  $\eta = h^*/h = 40$  produces stable results that are closely matched by the results obtained with  $\eta = 80$  (hence convergence through grid reduction is achieved).  $\eta = 20$  provides less satisfactory resolution. As an illustration, consider Figure 8, where the slip rate history during a part of the second event of the sequence with  $h^* = 0.94$  km is shown for the point at 3 km depth. Although in a different problem the value of  $\eta$  required may

be different, the above consideration illustrates how to approach selection of  $\eta = h^*/h$ .

We use  $\beta_{\text{min}}$  of 1/2, giving  $\Delta t_{\text{min}} = \Delta t_{\text{cell}}/2$ . For the convolution truncation, we choose the elastodynamic time windows by specifying  $T_w(1) = 2T_\lambda$  or  $T_w(1) = T_\lambda$  with  $T_\lambda = \lambda/c$ , and  $q_w = 4$ , confirming these values by varying the parameters to make sure the results do not depend on their choice. Note that since the replication period in our model is much larger than the region failing dynamically,  $T_w(1) = 2T_\lambda$  is quite a large window for this model, and almost identical results can be obtained by using  $T_w(1) = T_\lambda$ . Larger values of the elastodynamic time windows in this model tend to increase the maximum velocities achieved, by a small amount, but do not noticeably influence other features, such as the amount and distribution of slip or the rupture velocities.

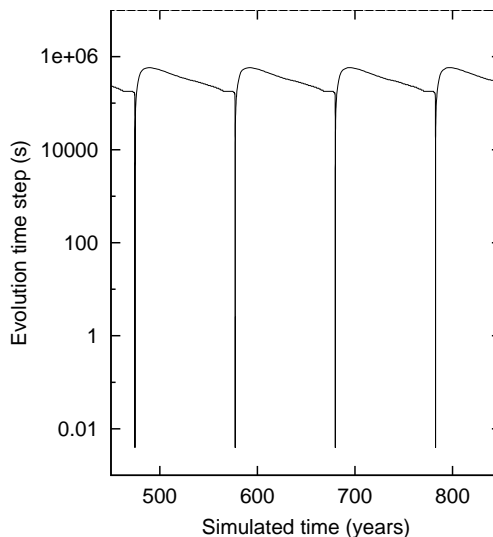
The fact that  $q_w$  can be chosen as small as 4 in this problem provides a huge computational gain. Let us illustrate that for our example with  $h^* = 0.94$  km. To achieve  $\eta = h^*/h = 40$ , we need  $N_{\text{ele}} = 8192$  elements along the replication period. This is also the number of Fourier modes for which we have to compute the convolution integrals at each time step (and hence to store slip velocity history sufficiently back in time). For mode  $n = 1$ , with  $T_w(1) = 2T_\lambda$



**Figure 8.** Slip velocity history at 3 km depth as a function of time for the second event in the sequence with  $h^* = 0.94$  km. Zero time is chosen arbitrarily for plotting convenience. The resolution  $h^*/h = 40$  gives essentially the same results as  $h^*/h = 80$ . The case  $h^*/h = 20$  is less sufficiently resolved.

and  $\beta_{\min} = 1/2$ , the number of values stored at spacing of  $\Delta t_{\min}$  is  $N_{\text{ele}}T_w(1)/(T_\lambda\beta_{\min}) = 4N_{\text{ele}} = 32,768$ . If a constant window were used for all modes, this would be the number of required values for each Fourier mode, and we would need two arrays of the size  $[32,768 \times 8192]$ , one for the Fourier coefficients and the other for the values of the discretized kernel. At each time step we would have to use these arrays to compute the convolution integrals. With  $q_w = 4$  and mode-dependent time windows, the number of values needed to be stored and used at each time step decreases significantly for higher Fourier modes. We pack more than one mode into most columns of the arrays, getting, for this particular example, the size of the arrays down to  $[32,768 \times 26]$ , which results in a very substantial reduction, by more than a factor of 300, of memory and cpu time for doing convolutions. This is consistent with our order-of-magnitude considerations in section 6.

Our evolution time step selection follows criterion (11) and (15)-(17). With the choice of other parameters discussed, the minimum evolution time step allowed,  $\Delta t_{\min}$ , is  $\approx 0.004$  s for the case  $h^* = 0.94$  km and  $\approx 0.001$  s for the case  $h^* = 0.235$  km. We can understand why such small values are needed by recognizing, as we did establishing criterion (11), that slip in one time step must be comparable (and preferably smaller) than the characteristic slip distance  $L$  of the friction law, to resolve the state variable evolution. Since the simulated  $L$  is of order millimeters (Figure 4a) and



**Figure 9.** Values of evolution time steps (in seconds) plotted as a function of the simulated time in years for the case  $h^* = 0.94$  km. Variable time stepping works well, making the time step span more than 8 orders of magnitude in this simulation, consistently with the changes in the slip velocity.

the largest slip velocities are of order meters per second, it is clear that the smallest time steps should be of the order of one thousandth of a second, as we have here. Note that once the dynamic propagation begins, large slip velocities at the rupture fronts determine the time step size, and it is almost always equal to the minimum time step  $\Delta t_{\min}$ , so that the rapid event propagation is essentially modeled using constant time steps. Preventing the time step from becoming smaller than  $\Delta t_{\min}$  may also mean that the state variable evolution is occasionally not ideally resolved right at the rupture peak, but in all the cases we have checked, not resolving  $L$  only occasionally at the very tip of the rupture does not change the results in any significant way.

For the slow deformation periods in between dynamic rupture events, the time steps taken are quite large. Figure 9 shows the values of evolution time steps for the case  $h^* = 0.94$  km. We see that the time steps span more than eight orders of magnitude in this simulation. Conditions (12) or (14) do a very good job restricting the time steps during slow deformation. If they are violated, the computation becomes corrupted. We can show this by relaxing the conditions several times, that is, by selecting the time step using (11) and (15) - (17) as before, but with coefficients  $\xi_i$  increased by a certain factor, although still insisting that  $\xi_i$  is not larger than  $1/2$ . Figure 10a shows the maximum velocity for the case  $h^* = 0.94$  km and a factor of 2 increase, and Figure 10b for a factor of 5 increase, of the time steps in the sense discussed above. Comparing to Figure 7a, which shows the well-resolved response, we see that numerical instabilities start to appear at slow sliding velocities for modestly increased time steps (Figure 10a), while for the factor of 5 increase the response looks very complex, with numerous events of different maximum velocities (Figure 10b), all of which are artifacts of the improper time discretization. Figure 11 shows slip accumulation for a factor of 3 increase, with some small "events" appearing (like the one at  $\approx 19$  m slip) and aperiodic large events. Those features are caused by the improper resolution in time; the true response in this case is the periodic sequence of large events shown in Figure 5. Improper space discretization also produces artificial complexification of the model response, as discussed by Rice [1993] and Ben-Zion and Rice [1995, 1997].

## 8. Discussion

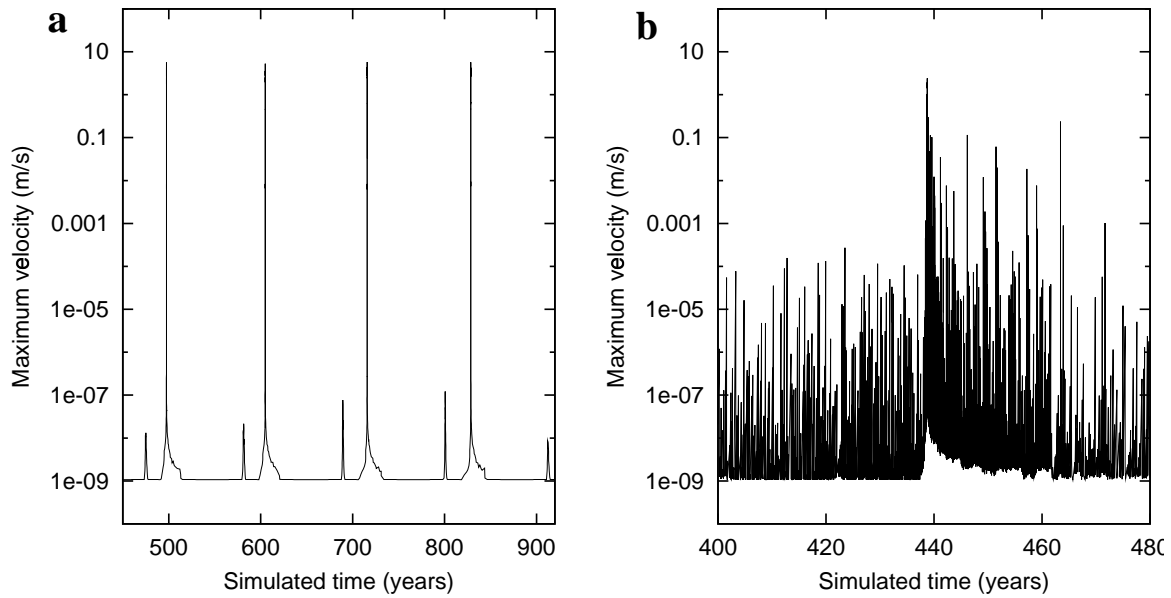
As the considered examples show, the algorithm presented here is capable of rigorous treatment of long-duration deformation histories with continuing aseismic creep slippage in velocity-strengthening fault regions throughout the loading period, with gradual nucleation of model earthquakes followed by dynamic propagation of ruptures, and

with rapid post seismic deformation after such events. The algorithm is formulated for general rate- and state-dependent friction laws, and the positive direct effect observed experimentally and represented by such laws is decisive for its success during long intervals with essentially quasi-static response and aseismic slip.

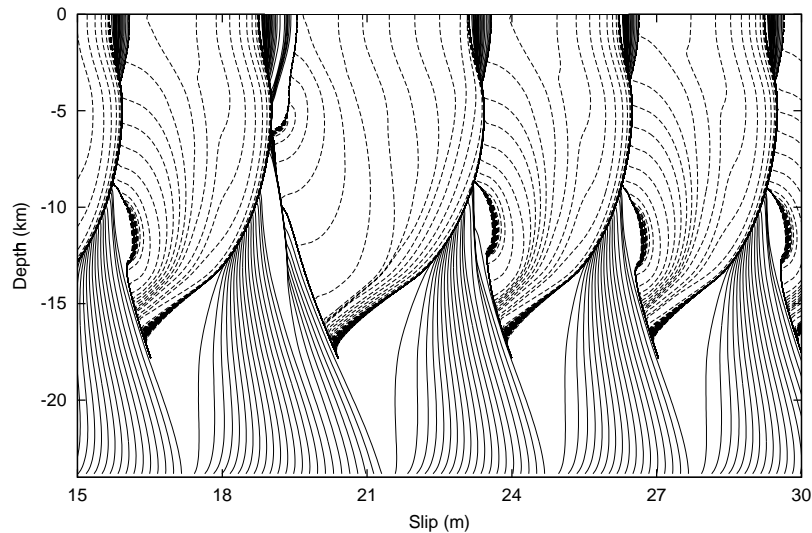
The algorithm employs a number of important ideas. Separation of the stress transfer functional into static and dynamic parts localizes the effects of the prior deformation history in convolution integrals on slip velocity with rapidly decaying kernels. Truncation of these convolutions is justified by rapid decay in time of the convolution kernels and allows us to simulate long processes without the necessity to deal with all prior deformation history at each time step. Variable time stepping makes the number of time steps during slow deformation periods numerically manageable while still capturing the details of both the nucleation and dynamic propagation phases. Proper space and time discretization ensures reliability of the results which can be verified through space and time grid refinement. The methodology has been presented using the 2-D antiplane spectral formulation. The described procedures can be readily extended to the 2-D in-plane and 3-D spectral formulations. They can also be applied at least in part to discretized models based on space-time boundary integral formulations, as we briefly discuss in Appendix B, where we further suggest that such formulations could be founded on kinematic modeling input from more versatile methods like finite difference, possibly being practical in cases for which the spectral diagonalization of convolutions does not apply.

The numerically most challenging parts of the algorithm are calculations of the convolution integrals and, in the spectral formulations, fast Fourier transforms (FFTs). If the elastodynamic time windows used to truncate the convolutions are long and truncation according to the mode does not shorten much the time windows for higher modes, then the convolution evaluations take most of the computational time and even minor optimizations of convolution evaluations can be very beneficial. If, on the other hand, the time windows are much shorter for higher modes, as we have been able to use for our depth-variable example here, then the FFTs start to use a comparable fraction of the computational time, and an efficient FFT routine can make a significant difference.

It is important to ensure that the results of the simulation do not depend on the discretization and other numerical parameters. For example, we could conclude that the model response is complex (Figure 10b) or that the events are aperiodic (Figure 11), whereas better resolution in time leads, in this case, to a periodic sequence of large events (Figures 5 and 7a). The verification of the independence of the results on numerics can be done through establishing convergence



**Figure 10.** Maximum velocity on the fault for  $h^* = 0.94$  km and insufficient resolution in time. (a)  $\xi_i$  increased by a factor of 2, (b)  $\xi_i$  increased by a factor of 5 (in both cases,  $\xi_i$  is not allowed to exceed  $1/2$ ). In Figure 10a, numerical instabilities start to show; Figure 10b shows seemingly very "complex" behavior, which is actually just a numerical artifact.



**Figure 11.** Accumulation of slip versus depth for the case  $h^* = 0.94$  km with insufficient resolution in time ( $\xi_i$  increased by a factor of 3). The solid lines are plotted every 5 years. The dashed lines are plotted above 18 km depth every second if the maximum velocity anywhere on the fault exceeds 0.001 m/s. The response differs from that of the properly resolved run in Figure 5. Nonperiodic large events and even some smaller events appear, all of which are, in this case, artifacts of the inadequate time discretization.

of the results (or at least of their qualitative features) as the parameters of the simulation are refined. This is often necessary even when the output looks smooth and plausible, as it can still be qualitatively different from the true response of the model (e.g., Figure 11 versus Figure 5). Usually, the plots of slip velocity or stress reveal much more about the numerical stability and convergence than their slip counterparts.

The examples presented are based on a rather simple model, with the fault properties uniform throughout the seismogenic (steady-state velocity-weakening) zone. The response consists of periodic sequences of events (Figures 5 and 6), and such a regular response has allowed us to concentrate on developing a rigorous and efficient numerical procedure. We verify, under the conditions of much better resolution and wider parameter range, the result of *Rice and Ben-Zion* [1996] and *Ben-Zion and Rice* [1997] that the dynamic effects alone are not sufficient to produce event complexity. Moreover, we find periodic response in some cases where the earlier studies have found, evidently due to insufficient numerical resolution, chaotic sequences of large events.

In our consideration, the characteristic slip distance  $L$  of the rate- and state-dependent friction is taken to be much larger than the laboratory values to achieve numerical tractability. To predict the model behavior with the laboratory-derived values of  $L$ , it is important to observe trends as we decrease  $L$ . Comparing Figures 5 and 6, with  $L$  in the case of Figure 6 being 4 times smaller than that of Figure 5, we see that the reduction in the characteristic slip distance introduced small events at the base of the seismogenic zone. Further twofold reduction in  $L$  produces a sequence of one large and one small event much like Figure 6 (although with smaller slip per event), but it is possible that still much smaller values of  $L$  would introduce more elaborate sequences, with more small events.

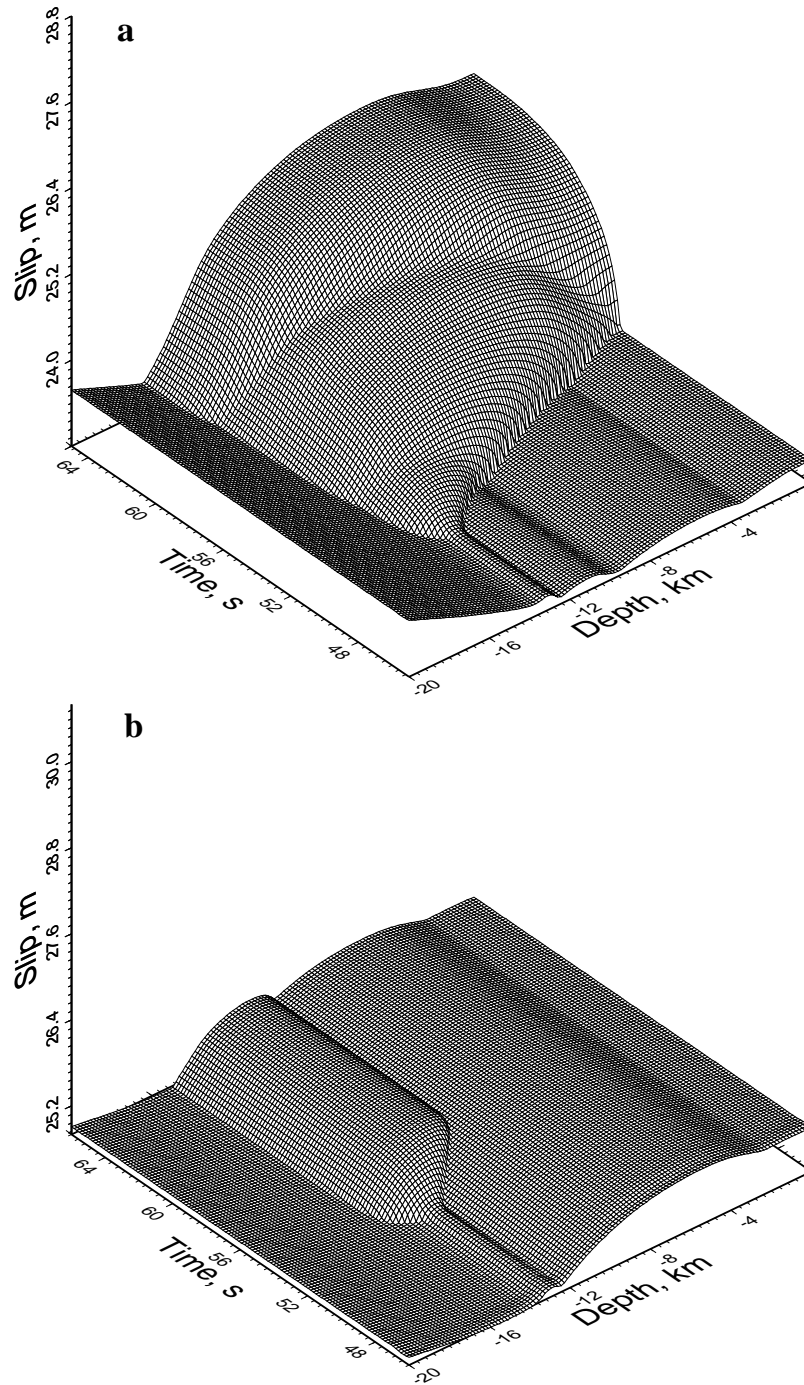
To produce realistically complex behavior, additional features have to be introduced. We would expect that adding strong fault heterogeneities in the form of (highly) nonuniform normal stress and/or frictional properties would naturally complexify the model response. Accounting for shear heating is also very important. It would introduce pore pressure development, which would add a second weakening mechanism due to the evolution of the effective normal stress. Interaction of two weakening mechanisms has complexified events sequences for a model studied by *Shaw and Rice* [2000], in a certain parameter range for the ad hoc type of friction law with two slip-weakening distances used. Another consequence of the shear heating would be temperature-induced time variations in the frictional properties, which may also contribute to event complexity. These problems, as well as other important problems such as the

earthquake nucleation process or patterns of rupture propagation in events nucleated naturally as a part of a sequence, can be studied within the methodology presented in this paper.

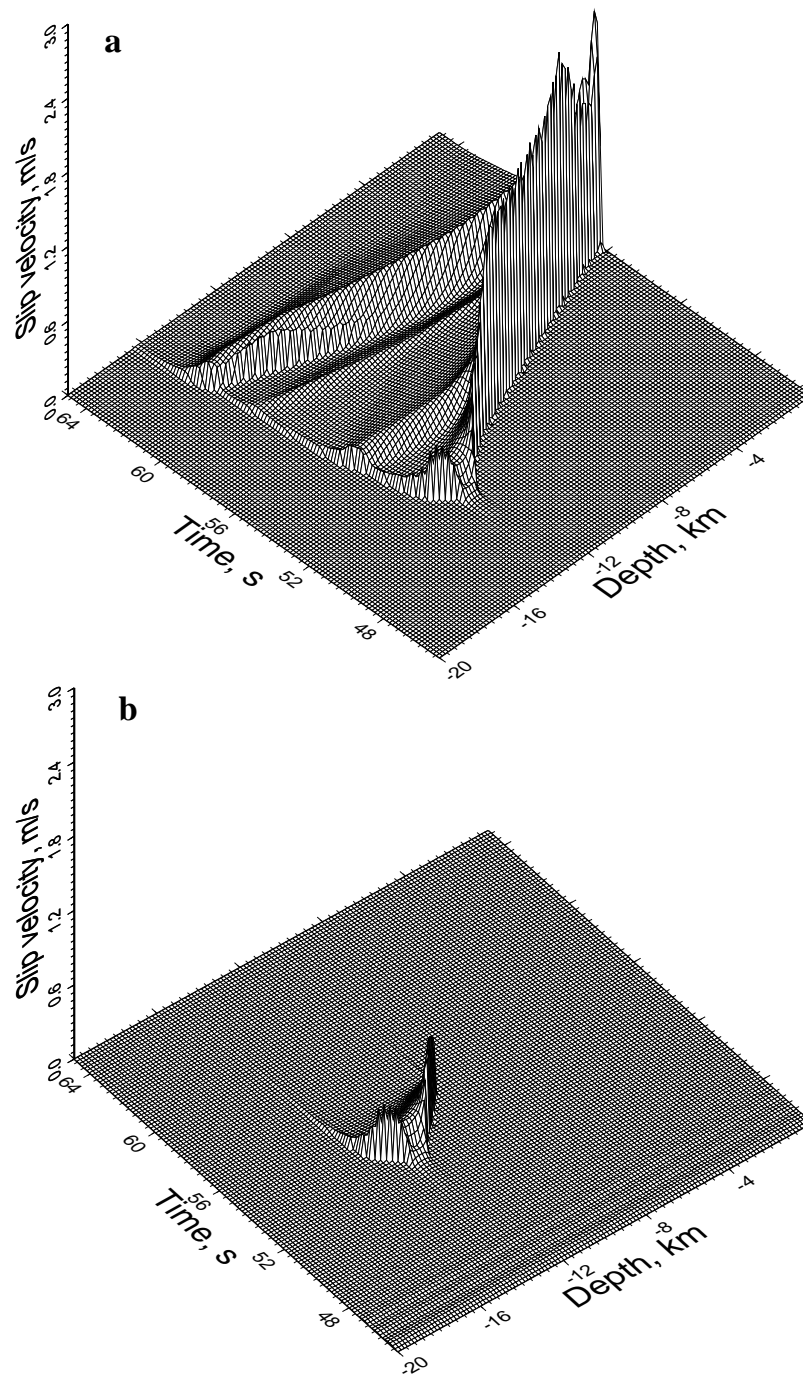
Comparison of the larger and the smaller events for the case  $h^* = 0.235$  km supports the view that large events are just small events that run away and shows how the runaway can be prevented by prior stress release. Let us consider the 3-D plots of slip and slip velocity for individual events shown in Figures 12 and 13. The distribution of slip before each of the shown model earthquakes (Figure 12) reflects the slip in previous events, the creeping velocity-strengthening regions on both ends of the fault segment, and a clear nucleation zone, which actually extends long back in time. The slip velocity plots (Figure 13) show how the dynamic events develop. Once an event nucleates, two rupture fronts propagate in the opposite directions. One of them is arrested in the velocity-strengthening region at the bottom of the fault. In the case of the larger event (Figure 13a), the other rupture front reflects off the free surface and runs down, rerupturing the seismogenic depth, with dynamic waves of slip propagating on the surface of the rupture. The spikes on the rupture front are an artifact of the outputting and plotting procedure; for a given space location as a function of time and for a given moment in time as a function of space, the slip velocity profiles are smooth. The plotting procedure also reduces the maximum slip velocities achieved, owing to insufficient resolution of the image surface. In the case of the smaller event (Figure 13b) the rupture gets arrested long before it reaches the free surface. From the slip distribution in Figure 12b we notice that the smaller event fails to advance into the region of larger slip (and hence higher stress release) left by the previous (larger) event. Comparing Figures 12a and 12b, as well as Figures 13a and 13b, we notice that the slips and slip velocities during and right after the nucleation of the smaller event look just like the ones for the beginning of the larger event. This means that observing signals from the nucleation and beginning of such an event, we would not be able to tell whether the final size of the event will be large or small.

We can use the developed methodology, which incorporates both truly slow tectonic loading and all dynamic effects, to evaluate simplified approaches. Let us consider two such approaches: (I) a procedure with truly slow tectonic loading but with a part of the dynamic effects (namely, dynamic stress transfers) ignored, and (II) a procedure with all dynamic effects incorporated but much faster loading.

To get a procedure of type I, we take  $T_w(1) = T_w = 0$ , which coincides with the quasi-dynamic approximation used by *Rice* [1993] and *Ben-Zion and Rice* [1995]. Figure 14 shows the results for the case  $h^* = 0.94$  km. Comparing

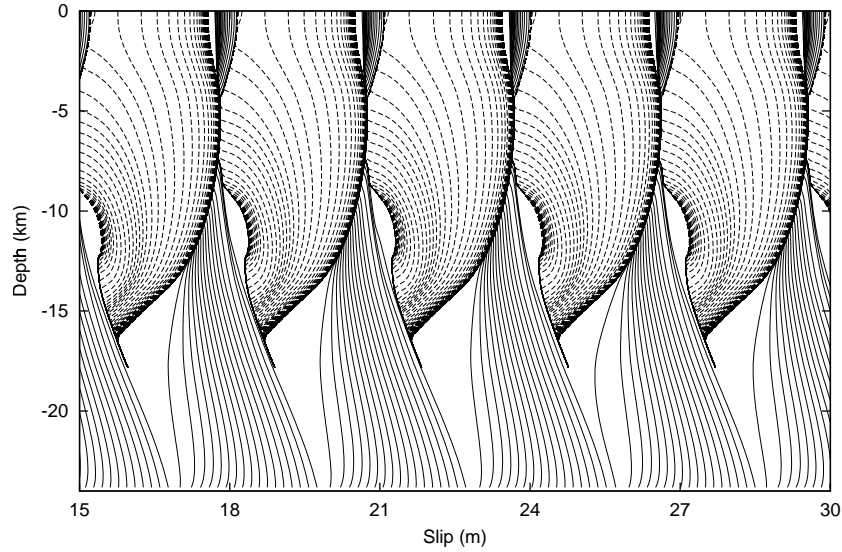


**Figure 12.** Slip in individual events from the sequence in Figure 6,  $h^* = 0.235$  km, for (a) a larger event and (b) the following smaller event. The time axis spans 20 seconds, with zero time chosen arbitrarily for plotting convenience. The slip axis spans 6 m in both cases. Notice the clear nucleation zone that extends much further back in time. The smaller event in Figure 12b looks just like the beginning of the larger event in Figure 12a; it stops by not being able to advance into the higher slip/lower stress region in the middle of the fault. This supports the idea that large events are small events that run away due to favorable stress/strength conditions on the fault.

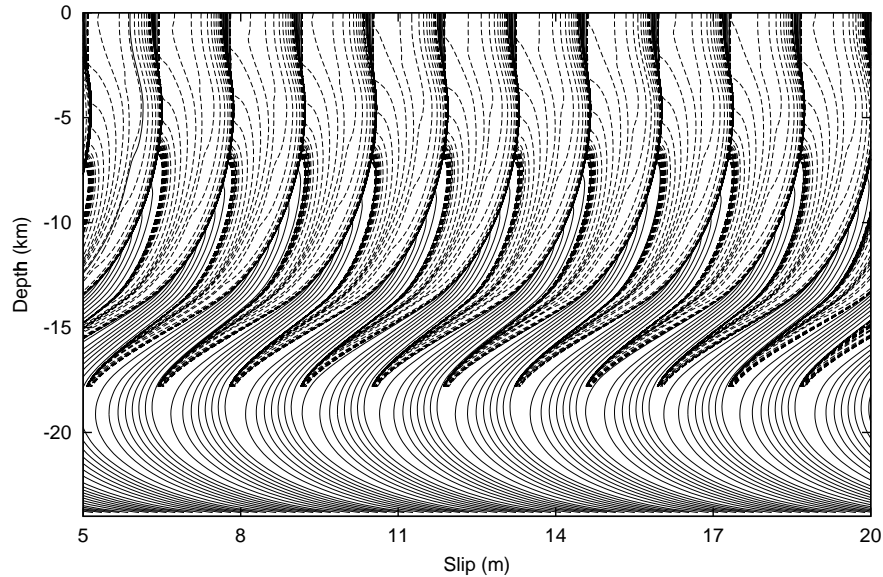


**Figure 13.** Slip velocity in the same events as in Figure 8,  $h^* = 0.235$  km, for (a) a larger event and (b) the following smaller event. The spikes on the rupture front are artifacts of the outputting and plotting procedures; the slip velocity is actually smooth, as confirmed by plots of slip velocity at particular depths as a function of time and at particular times as a function of depth. Notice the dynamic effects in the form of waves, including the one reflected from the free surface. As in Figure 12, the smaller event in Figure 13b looks just like the beginning of the larger event in Figure 13a.





**Figure 14.** Accumulation of slip versus depth for the case  $h^* = 0.94$  km with  $T_w = 0$  (i.e., the "quasi-dynamic" approximation). The solid lines are plotted every 5 years. The dashed lines are plotted above 18 km depth every second if the maximum velocity anywhere on the fault exceeds 0.001 m/s. There is no reflected front of slip from the free surface which is present when the wave-mediated stress transfers are included (Figure 5). The slip per event and the slip and rupture velocities are smaller.



**Figure 15.** Accumulation of slip versus depth for the case  $h^* = 0.94$  km with the much larger loading velocity  $V_{pl} = 35 \times 10^6$  mm/year. The solid lines are plotted every  $5 \times 10^{-6}$  years. The dashed lines are plotted above 18 km depth every second if the maximum velocity anywhere on the fault exceeds 0.001 m/s. (The initial conditions could be tuned so that the slip in the middle of the lower velocity strengthening region is roughly the same as at its bottom. The qualitative features of the simulation are independent of the initial conditions.) Notice that, compared to the case of tectonic loading ( $V_{pl} = 35$  m/s, Figure 5), the nucleation phase is very different and the slip per event is more than twice smaller.

with Figure 5, we see that the wave effects disappear as they should (there is no slip wave reflecting off the free surface), the rupture velocities are slower (the tips of the dashed lines are closer to one another), and the slip velocities are smaller too, as is the accumulated slip per event. We have examined the possibility that the quasi-dynamic calculations ( $T_w = 0$ ) might approximately reproduce the slip per event as in Figure 5 if we could make the slip velocities faster during dynamic events. To that end, we did a series of studies in which the numerical value of  $\mu/2c$  in the radiation damping term of elastodynamic relation (26) was varied in size. Reduction to approximately half the proper value did give rupture propagation and slip velocities during events that tend to be of comparable order to those of the proper dynamic simulation (Figure 5), but there was very little effect on the slip per cycle, which remained much as in Figure 14. This suggests that the greater slip in Figure 5 is significantly due to the discussed reflected wave effect at the free surface, a feature which could not appear in the quasi-dynamic simulation. However, there is no qualitative difference in the system behavior, probably due to the relative simplicity of this model's response. If the dynamic effects were more complex, their elimination may have made a more significant difference.

Finally, we consider a procedure in which the plate loading rate is only a few orders of magnitude less than representative seismic slip rates, which allows to use standard elastodynamic numerical methodology throughout (like in the work by *Shaw and Rice* [2000]). To this end, we change the loading velocity from  $V_{pl} = 35$  mm/yr in our implementation example above to  $V_{pl} = 35 \times 10^6$  mm/yr, for the case  $h^* = 0.94$  km, keeping other parameters the same. The resultant slip accumulation is shown in Figure 15, which is very different from the model response with the tectonic-like loading velocity (Figure 5). The increase in loading velocity has totally altered the nucleation process and location and resulted in more than twice smaller slip per event ( $\approx 1.4$  m slip per event, compared to  $\approx 3.6$  m slip per event in the case of  $V_{pl} = 35$  mm/yr). The loading velocities  $V_{pl} = 35 \times 10^5$  mm/yr and  $V_{pl} = 35 \times 10^4$  mm/yr also give quite altered (although, naturally, less so) nucleation, and slips per event of  $\approx 1.8$  and 2.3 m, respectively.

The evaluation of the simplified approaches shows that while accounting for the proper dynamic response can be very important, simulating slow tectonic loading is also crucial for uncovering the true model response. The results also demonstrate that even though simplified approaches may be unavoidable in some cases, they have limitations that can be uncovered and remedied only within a more general approach like the one presented here.

## Appendix A: Derivation of Critical Stiffness and Time Discretization Constraints Dictated by Constitutive Law

To derive expression (9) for the critical stiffness  $k_{cr}$  in quasi-static slip, let us consider a spring-slider system moving steadily at rate  $V^*$ , with state  $\theta^* = \theta_{ss}(V^*)$  and frictional resistance  $\tau^* = \tau_{ss}(V^*)$ . If we impose a perturbation  $\tilde{V}$  on the sliding velocity and use  $\tilde{\tau}$ ,  $\tilde{\theta}$ , and  $\tilde{\delta}$  to denote the corresponding perturbations of other quantities, then from the friction law (7) we can write the linearized constitutive response to the perturbation as

$$\tilde{\tau} = A^* \frac{\tilde{V}}{V^*} + C^* \tilde{\theta}, \quad (\text{A1a})$$

$$C^* \dot{\tilde{\theta}} = -\frac{V^*}{L} \left( C^* \tilde{\theta} + \frac{B^*}{V^*} \tilde{V} \right),$$

where  $A^*$ ,  $B^*$ ,  $C^*$ , and  $L$  are used to denote the partial derivatives of the functions in (7):  $\partial\psi(V, \theta)/\partial V = A^*/V^*$ ,  $\partial\psi(V, \theta)/\partial\theta = C^*$ ,  $\partial\varphi(V, \theta)/\partial V = -B^*/LC^*$ , with the derivatives evaluated at steady state, and  $L$  given by (9b). The definitions of  $A^*$  and  $B^*$  are equivalent to the more familiar  $A^* = [V \partial\tau(V, \theta)/\partial V]_{V=V^*, \theta=\theta_{ss}(V^*)}$ ,  $A^* - B^* = [V d\tau_{ss}(V)/dV]_{V=V^*}$ .

At steady state ( $\dot{\tilde{\theta}} = 0$ ) we have  $\tilde{\tau} = (A^* - B^*)\tilde{V}/V^*$ , and thus the sign of  $(B^* - A^*)$  determines whether the system exhibits steady-state velocity weakening or strengthening.  $A^*$  and  $B^*$  can, in general, depend on the steady-state velocity  $V^*$ , but in the most commonly used Dieterich-Ruina logarithmic forms of the rate- and state-dependent friction laws, they are chosen as constants times effective normal stress (as we do in our implementation example) and are often denoted by  $A$  and  $B$ . The (quasi-static) elastic response to the perturbation is given by

$$\tilde{\tau} = -k\tilde{\delta}, \quad (\text{A1b})$$

where  $k$  is the spring stiffness. Eliminating  $\tilde{\tau}$  by combining (A1a) and (A1b), substituting  $\tilde{V} = \tilde{\delta}$ , and defining, for simplification, nondimensional quantities  $D = \tilde{\delta}/L$ ,  $\Theta = C^*\tilde{\theta}/B^*$ , and  $T = V^*t/L$ , we obtain the following simple system of two linear differential equations:

$$\left\{ \begin{array}{l} dD/dT \\ d\Theta/dT \end{array} \right\} = M \left\{ \begin{array}{l} D \\ \Theta \end{array} \right\}, \quad (\text{A2a})$$

$$M = \begin{bmatrix} -kL/A^* & -B^*/A^* \\ kL/A^* & (B^* - A^*)/A^* \end{bmatrix}. \quad (\text{A2b})$$

Qualitative behavior of such systems depends on the eigenvalues of the matrix  $M$ . If the real parts of all the eigenvalues are negative, the solution vanishes with increasing time, and if at least one eigenvalue has a positive real part, the solution grows without limit. When the largest real part is zero, long-time periodic motion results. In the system (A2), the perturbation grows (exponentially) for  $\rho < 0$  and decays (exponentially) for  $\rho > 0$ , where

$$\rho = \frac{L}{A^*} \left[ k - \frac{B^* - A^*}{L} \right]. \quad (\text{A3})$$

For velocity strengthening friction,  $(B^* - A^*) < 0$ , we have  $\rho > 0$  and the perturbation always decays. For velocity weakening friction,  $(B^* - A^*) > 0$ , the behavior depends on the sign of the expression in brackets of (A3), which defines the critical stiffness

$$k_{\text{cr}} = \frac{B^* - A^*}{L} = \left[ -\frac{V d\tau_{\text{ss}}(V)/dV}{L} \right]_{V=V^*}, \quad (\text{A4})$$

and the perturbation grows in time for  $k < k_{\text{cr}}$  and decays in time for  $k > k_{\text{cr}}$ , as claimed in (9a).

To derive the constraint (12) on the evolution time stepping, we analyze, as a simple model case, explicit integration (with a constant time step  $\Delta T = V^* \Delta t / L$ ) of the governing system for the perturbation (A2). In the system (A2), inertia effects are ignored, but they are negligible at low slip rates. The discretized system is given by

$$\begin{Bmatrix} (D_{n+1} - D_n)/\Delta T \\ (\Theta_{n+1} - \Theta_n)/\Delta T \end{Bmatrix} = M \begin{Bmatrix} D_n \\ \Theta_n \end{Bmatrix}, \quad (\text{A5a})$$

or

$$\begin{Bmatrix} D_{n+1} \\ \Theta_{n+1} \end{Bmatrix} = Q \begin{Bmatrix} D_n \\ \Theta_n \end{Bmatrix}, \quad (\text{A5b})$$

with

$$Q = \begin{bmatrix} 1 - kL\Delta T/A^* & -B^*\Delta T/A^* \\ kL\Delta T/A^* & 1 + (B^* - A^*)\Delta T/A^* \end{bmatrix}. \quad (\text{A5c})$$

We would like this discrete dynamic system to have the same stability behavior as the continuous (in time) system (A2) does. The behavior of such discrete dynamic systems as (A5) can be understood by finding the eigenvalues of the governing matrix  $Q$  and comparing their absolute values to unity. The eigenvalues of matrix  $Q$  are given by

$$\lambda_{1,2} = 1 + \left( -\frac{1}{2}\rho \pm \sqrt{\frac{1}{4}\rho^2 - \frac{kL}{A^*}} \right) \Delta T, \quad (\text{A6})$$

with  $\rho$  defined by (A3).

Considering eigenvalues (A6), we conclude the following. If  $\rho < 0$ , which corresponds to the situation with steady-state velocity weakening  $(B^* - A^*) > 0$  and improperly sparse grid (giving  $k - (B^* - A^*)/L = k - k_{\text{cr}} < 0$ ), the eigenvalues have absolute values larger than unity, and the perturbation grows, consistently with the continuous case, regardless of the chosen time step  $\Delta T$ . For  $\rho > 0$ , which corresponds to either steady-state velocity strengthening, or steady-state velocity weakening with a sufficiently dense grid, the continuous case predicts decay of the perturbation. Indeed, the discretized case follows this stable behavior if the time step  $\Delta T = V^* \Delta t / L$  is sufficiently small. That is, in the case  $\rho > 0$ , the absolute values of the eigenvalues is less than unity, and hence the perturbation decays, if the (dimensional) time step  $\Delta t$  satisfies the conditions written in the text as equations (12).

## Appendix B: Remarks on Extension of Methodology to Cases With No Translational Invariance

Our methodology can be extended to general cases which lack translational invariance, to accommodate problems such as a fault oblique to the free surface or heterogeneity of bulk material properties (like a layered Earth structure). If a boundary integral equation is discretized with  $N_{\text{sp}}$  sample points (nodes) over the fault domain, the stress transfer functional will still be related to the slip by an expression which can be put in the velocity form

$$f_i(t) = - \sum_{j=1}^{N_{\text{sp}}} [K_{ij} \delta_j(t) - \int_0^t C_{ij}(t-t') \dot{\delta}_j(t') dt'], \quad (\text{B1})$$

where subscripts  $i$  and  $j$  denote values discretized in space. Here  $[K]$  is the static elastic stiffness matrix, and  $[C(t)]$  is a matrix of convolution kernels. These could emerge as the result of discretizing a boundary integral formulation, or could be determined by standard elastodynamic finite difference or finite element calculations. The calculations would be purely kinematic, imposing a step in slip at each node singly (say at node  $j$ ), over a single elastodynamic time step, and calculating the stress histories  $\tau_i(t)$  at all other nodes  $i$ . By analyzing the  $\tau_i(t)$  using (1) and (B1), the  $K_{ij}$  could be extracted numerically as the long-term limit of the stress changes and the  $C_{ij}(t)$  determined from the transient response and tabulated.

The relation (B1) is analogous to (2) and (4), and most ideas of the presented algorithm may be directly applied, including the truncation of convolutions, the schemes to advance through a time step, and the time step selection pro-

cedure. One exception is the notion of shorter truncation windows for higher mode numbers, which naturally arises in the spectral formulation and greatly reduces storage requirements and execution time compared to a uniform truncation window for all modes, as discussed in section 6. Unfortunately, there is no directly similar notion for the general case (B1).

The following concept may, however, provide the extension of the spectral formulation to the general case (B1), enabling the possibility of shorter truncation windows for some time convolutions in (B1). We expect that even in the absence of translational invariance, if we read into (B1) a step-in-time slip distribution which varies spatially like  $\text{Re}(e^{ikz})$ , or like  $\text{Re}(e^{i(kz+mx)})$  in 3-D modeling of a fault which spans two space dimensions  $x$  and  $z$ , then for high  $|k|$ , or  $\sqrt{k^2 + m^2}$ , there would generally be rapid decay of the transient stress in approach to the static limit. In such a case, the convolution could be truncated earlier than for lower  $|k|$  or  $\sqrt{k^2 + m^2}$ . Hence, the following procedure can be explored, illustrated further for the 2-D case. The  $N_{\text{sp}}$  slips  $\delta_i(t)$  can be represented as a linear combination of  $N_{\text{sp}}$  independent basis functions, each with progressively higher wave number features. While some wavelet expansion may turn out to be a preferred way of accomplishing that, a simple illustration is provided by using the Fourier sum like in (3) for the  $\delta_j(t)$  and  $f_j(t)$ . Using the  $R$  notation of section 5 for this representation and its inversion, with  $\delta_j = \sum_m R_{jm}^{-1} D_m$  and  $F_n = \sum_i R_{ni} f_i$ , we can transform the relation (B1) to

$$F_n(t) = - \sum_m [\hat{K}_{nm} D_m(t) - \int_0^t \hat{C}_{nm}(t-t') \dot{D}_m(t') dt'], \quad (\text{B2})$$

where

$$[\hat{K}_{nm}, \hat{C}_{nm}(t)] = \sum_i \sum_j R_{ni} [K_{ij}, C_{ij}(t)] R_{jm}^{-1},$$

summations on  $i$  and  $j$  extend over  $[1, N_{\text{sp}}]$  and on  $m$  over  $[-N_{\text{sp}}/2, N_{\text{sp}}/2]$ . One may hope that truncation times  $T_w(n, m)$  for  $\hat{C}_{nm}(t)$  can be chosen shorter for the higher wave numbers represented by, say, the maximum of  $|n|$  or  $|m|$ .

In implementing such an approach, one can determine  $\hat{K}_{nm}$  and  $\hat{C}_{nm}(t)$  directly, without determining  $K_{nm}$  and  $C_{nm}(t)$  first. The response to the slip history  $\delta_j$ , which is a step function in time and Fourier mode  $m$  (or mode  $m$  of the wavelet expansion used) in space, would be determined, and hence values corresponding to  $f_i(t)$  would be known. Then

fast transforms would be done on these computed values, to convert them to histories  $F_n(t)$ . From (B2), components of  $\hat{K}_{nm}$  and  $\hat{C}_{nm}(t)$  would be given as  $\hat{K}_{nm} = -F_n(\infty)$  and  $\hat{C}_{nm}(t) = F_n(t) - F_n(\infty)$ , for the value of  $m$  used in the slip history selected. The procedure would have to be repeated for all values of  $m$ .

For cases without translational invariance, there is no simple transformation to diagonalize the matrix  $[C(t)]$  from (B1), and the matrix  $[\hat{C}(t)]$  from (B2) will generally not be diagonal. For the problems that have translational invariance (or can be mapped to such by addition of a mirror image), the matrix  $[C(t)]$  in the space-time formulation (B1) is still nondiagonal, even though it acquires a special form with  $C_{ij} = C_{|i-j|}$ . However, in this case the spectral formulation considered in section 2 diagonalizes the matrix of the convolution kernels. This translates into substantial reduction in the number of time convolutions and makes the spectral formulation of a translationally invariant problem (such as (2) and (4)) computationally much more efficient than the space-time formulation of the same problem, or than any formulations of problems that lack translational invariance. This is true despite the additional cost of the spectral formulation, which needs more degrees of freedom. Suppose that the replication distance  $\lambda$  of the spectral formulation is  $q$  times larger than the domain which we would like to simulate (so that  $N_{\text{ele}} = qN_{\text{sp}}$ ), and that the time for computation of one truncated time convolution integral scales linearly with the number of space elements involved (as is true when all truncation windows are comparable to the time for the elastic waves to propagate through the simulated region). In a formulation with a nondiagonal matrix of convolution kernels, we need  $N_{\text{sp}}^2$  time convolutions to compute the values of the (discretized) functional  $f$  for all nodes. This requires the number of floating point operations proportional to  $N_{\text{sp}}^2 \times N_{\text{sp}} = N_{\text{sp}}^3$ . In the spectral formulation, such as (3) and (7), due to the diagonalization, we need only  $N_{\text{ele}} = qN_{\text{sp}}$  time convolutions, one per each Fourier mode. This translates into the number of operations proportional to  $q^2 N_{\text{sp}}^2$ . The FFTs used in the spectral formulation add the number of operations proportional to  $qN_{\text{sp}} \log_2(qN_{\text{sp}})$ , so that the overall number of operations in the spectral formulation is still proportional to  $q^2 N_{\text{sp}}^2$ . Since  $q$  is usually small number, such as 4 to 8, and typical values for  $N_{\text{sp}}$  are 512 to 65536, comparing  $N_{\text{sp}}^3$  and  $q^2 N_{\text{sp}}^2$  immediately shows that the spectral representation of a translationally invariant problem is computationally much more efficient.

**Acknowledgments.** These studies were supported by the Southern California Earthquake Center (SCEC), by USGS grant 99-HQ-GR-0025 to Harvard University, and by a Blaise Pascal International Research Chair of the Foundation of Ecole Normale Supérieure, Paris, to JRR. SCEC is supported by NSF Cooperative

Agreement EAR-8920136 and USGS Cooperative Agreements 14-08-0001-A0899 and 1434-HQ-97AG01718. This is SCEC contribution 511. We are grateful to John Morrissey for participation in earlier phases of this research and to Alain Cochard for extensive discussion and comments on the manuscript.

## References

- Baumberger, T., Contact dynamics and friction at a solid-solid interface: material versus statistical aspects, *Solid State Commun.*, *102*, 175-185, 1997.
- Ben-Zion, Y., and J. R. Rice, Slip patterns and earthquake populations along different classes of faults in elastic solids, *J. Geophys. Res.*, *100*, 12,959-12,983, 1995.
- Ben-Zion, Y., and J. R. Rice, Dynamic simulations of slip on a smooth fault in an elastic solid, *J. Geophys. Res.*, *102*, 17,771-17,784, 1997.
- Blanpied, M. L., D. A. Lockner, and J. D. Byerlee, Fault stability inferred from granite sliding experiments at hydrothermal conditions, *Geophys. Res. Lett.*, *18*(4), 609-612, 1991.
- Blanpied, M. L., D. A. Lockner, and J. D. Byerlee, Frictional slip of granite at hydrothermal conditions, *J. Geophys. Res.*, *100*, 13,045-13,064, 1995.
- Boatwright, J., and M. Cocco, Frictional constraints on crustal faulting, *J. Geophys. Res.*, *101*, 13,895-13,909, 1996.
- Brechet, Y., and Y. Estrin, The effect of strain rate sensitivity on dynamic friction of metals, *Scripta Metall. Mater.*, *30*, 1449-1454, 1994.
- Chester, F. M., Effects of temperature on friction: Constitutive equations and experiments with fault gouge, *J. Geophys. Res.*, *99*, 7247-7261, 1994.
- Chester, F. M., and N. Higgs, Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions, *J. Geophys. Res.*, *97*, 1859-1870, 1992.
- Cochard, A., and R. Madariaga, Dynamic faulting under rate-dependent friction, *Pure Appl. Geophys.*, *142*, 419-445, 1994.
- Cochard, A., and R. Madariaga, Complexity of seismicity due to highly rate dependent friction, *J. Geophys. Res.*, *101*, 25,321-25,336, 1996.
- Cochard, A., and J. R. Rice, A spectral method for numerical elastodynamic fracture analysis without spatial replication of the rupture event, *J. Mech. Phys. Solid*, *45*, 1393-1418, 1997.
- Dieterich, J. H., Modeling of rock friction, 1. Experimental results and constitutive equations, *J. Geophys. Res.*, *84*, 2161-2168, 1979.
- Dieterich, J. H., Constitutive properties of faults with simulated gouge, in *Mech. Behavior Crustal Rocks*, edited by N. L. Carter et al, *Geophys. Monogr. Ser.*, vol. 24, pp. 103-120, AGU, Washington, D.C., 1981.
- Dieterich, J. H., Earthquake nucleation on faults with rate- and state-dependent strength, *Tectonophysics*, *211*, 115-134, 1992.
- Dieterich, J. H., A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, *99*, 2061-2618, 1994.
- Geubelle, P. H., and J. R. Rice, A spectral method for 3D elastodynamic fracture problems, *J. Mech. Phys. Solids*, *43*, 1791-1824, 1995.
- Gu, J.-C., J. R. Rice, A. L. Ruina, and S. T. Tse, Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction, *J. Mech. Phys. Solids*, *32*, 167-196, 1984.
- Horowitz, F. G., and A. Ruina, Slip patterns in a spatially homogeneous fault model, *J. Geophys. Res.*, *94*, 10,279-10,298, 1989.
- Morrissey, J. W., and P. H. Geubelle, A numerical scheme for mode III dynamic fracture problems, *Int. J. Numer. Meth. Eng.*, *40*, 1181-1196, 1997.
- Myers, C. R., B. E. Shaw, and J. S. Langer, Slip complexity in a crustal-plane model of an earthquake fault, *Phys. Rev. Lett.*, *77*, 972-975, 1996.
- Okubo, P. G., Dynamic rupture modeling with laboratory-derived constitutive relations, *J. Geophys. Res.*, *94*, 12,321-12,335, 1989.
- Perrin, G., J. R. Rice, and G. Zheng, Self-healing slip pulse on a frictional surface, *J. Mech. Phys. Solids*, *43*, 1461-1495, 1995.
- Ranjith, K., and J. R. Rice, Stability of quasi-static slip in a single degree of freedom elastic system with rate and state dependent friction, *J. Mech. Phys. Solid*, *47*, 1207-1218, 1999.
- Rice, J. R., Spatio-temporal complexity of slip on a fault, *J. Geophys. Res.*, *98*, 9885-9907, 1993.
- Rice, J. R., and Y. Ben-Zion, Slip complexity in earthquake fault models, *Proc. Natl. Acad. Sci. U.S.A.*, *93*, 3811-3818, 1996.
- Rice, J. R., and A. L. Ruina, Stability of steady frictional slipping, *J. Appl. Mech.*, *50*, 343-349, 1983.
- Ruina, A. L., Slip instability and state variable friction laws, *J. Geophys. Res.*, *88*, 10,359-10,370, 1983.
- Shaw, B. E., and J. R. Rice, Existence of continuum complexity in the elastodynamics of repeated fault ruptures, *J. Geophys. Res.*, in press, 2000.
- Shibazaki, B., and M. Matsu'ura, Spontaneous processes for nucleation, dynamic propagation, and stop of earthquake rupture, *Geophys. Res. Lett.*, *19*, 1189-1192, 1992.
- Stuart, W. D., Forecast model for the great earthquakes at the Nankai Trough subduction zone, *Pure Appl. Geophys.*, *126*, 619-641, 1988.
- Stuart, W. D., and T. E. Tullis, Fault model for preseismic deformation at Parkfield, California, *J. Geophys. Res.*, *100*, 24,079-24,099, 1995.
- Tse, S. T., and J. R. Rice, Crustal earthquake instability in relation to the depth variation of frictional slip properties, *J. Geophys. Res.*, *91*, 9452-9472, 1986.
- Tullis T. E., Rock friction and its implications for earthquake prediction examined via models of Parkfield earthquakes, *Proc. Natl. Acad. Sci. U.S.A.*, *93*, 3803-3810, 1996.
- Tullis, T. E., and J. D. Weeks, Constitutive behavior and stability of frictional sliding of granite, *Pure Appl. Geophys.*, *124*, 383-414, 1986.
- Zheng, G., and J. R. Rice, Conditions under which velocity-weakening friction allows a self-healing versus a crack-like mode of rupture, *Bull. Seismol. Soc. Am.*, *88*, 1466-1483, 1998.
- Zheng, G., Y. Ben-Zion, J. R. Rice, and J. Morrissey, A new method for calculating elastodynamic response over long stressing histories containing rapid rupture episodes (abstract), *Eos Trans.*

*AGU*, 76(46), Fall Mtg. Suppl., F405, 1995.

---

N. Lapusta and J. R. Rice, Division of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138. (lapusta@esag.harvard.edu; rice@esag.harvard.edu)

Y. Ben-Zion, Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089-0740. (benzion@topaz.usc.edu)

G. Zheng, Technology Group, IBM, Somers, NY 10589. (gutuan@us.ibm.com)

Received October 18, 1999; revised May 31, 2000; accepted July 13, 2000.

---

This preprint was prepared with AGU's  $\LaTeX$  macros v5.01, with the extension package 'AGU++' by P. W. Daly, version 1.5e from 1997/11/18.