Notes on modeling earthquake dynamics
with the spectral element method

Jean-Paul Ampuero[*]

February 10, 2009

Abstract

We provide a detailed presentation and discussion of a numerical modeling technique for earthquake dynamics based on the spectral element method. The problem of dynamic frictional contact along slip interfaces in elastic domains is formulated in a hybrid variational form, with Lagrange multipliers associated to the fault tractions. The problem is discretized in space by a spectral element method and in time by a second order finite difference time scheme. We investigate the convergence of the numerical error with grid refinement in canonical test cases. We study in detail the origin and nature of the high frequency numerical noise as well as different solutions to reduce it. Comparisons with other methods on a standard convergence test show that the spectral element method for fault dynamics is very accurate and competitive. The ability of the method to handle rupture problems with realistic geometries, fault zone structure and non-linear off-fault rheology is illustrated through several examples.

[*]Institute of Geophysics, ETH Zürich, `ampuero@erdw.ethz.ch`

Contents

1 Introduction

1.1 Motivations

Three active areas of application of computational earthquake dynamics (CED) impose particularly challenging demands on numerical modeling methods. First, the study of fundamental problems in earthquake physics require the exploration of vast parameter spaces, to assess the general effect of a large number of uncertain assumptions. Second, the emerging role of earthquake dynamics in physics-based methodologies for earthquake hazard assessment requires dynamic source simulations with multiple realizations of stochastic, spatially non uniform input initial stresses and fault properties to evaluate the statistics of ground motion parameters [e.g. Ripperger et al 2007]. Third, there is a recent trend in earthquake seismology towards dynamic source inversion approaches, an inverse problem formulated in terms of the fault strength parameters and initial stresses. These applications typically require hundreds to thousands of large scale forward simulations.

Modern methods for CED should be able to include the following physical ingredients: free surface, non planar fault geometry (curved, rough, branched) and multiple fault interaction, heterogeneous crustal velocity model, multi-physics coupling in the fault zone (non-elastic bulk rheologies, thermal and fluid processes), multiple time-scales (seismic, aseismic and transitional). This requires a flexible computational framework.

The ratio of relevant length scales is very large ($> 10^4$). On the one hand, the smallest relevant scale for rupture dynamics (the size of the breakdown zone), the shortest wavelength at frequencies relevant for engineering (up to 20 Hz) and the scales of shallow sedimentary layering are < 100 m. On the other hand the total source dimensions and propagation distances are > 100 km. This places high demands on the performance (accuracy/cost trade-off) of numerical methods for wave propagation and source dynamics.

Brief overview of existing methods for CED and their drawbacks: FEM, FDM, BIEM, FVM. From outside the CED community: XFEM.

1.2 On singularities and discontinuities in rupture dynamic problems

Irregularities in the exact solution of a physical problem can affect the quality of its numerical solutions. In general the order of a physical singularity or discontinuity places limits on the order of convergence of the numerical error as a function of grid size [1]. Problems in rupture dynamics typically contain singularities and discontinuities arising from the fault constitutive law or from the initial and arrest conditions. These are reviewed in this section.

In ideal crack models a strong singularity appears at the leading edge of the rupture front: stress and velocities have an inverse square-root behavior as a function of the distance to the crack tip. This is usually regularized by the introduction of a cohesive zone (also called process zone or breakdown zone) with finite energy and finite strength [Dugdale; Barenblatt]. In CED, the most widely adopted approach for such regularization is the linear slip weakening friction law [Ida, 1972; Palmer and Rice, 1973; Andrews, 1975] [2], a first order representation of experimental observations of frictional weakening [Ohnaka; Dieterich]. Linear slip weakening guarantees continuity of velocity and fault shear stress but acceleration remains singular at the leading and trailing edges of the process zone [Ida, 1973]. The trailing edge singularity is

---

[1] add a classical reference here, or cite Karniadakis and Sherwin's book
[2] add figure here

readily avoided by assuming a smooth slip weakening curve close to the critical slip distance (D_c). Avoiding the singularity of acceleration at the leading edge requires an appropriate initial strengthening phase in the slip dependent friction law, but its design is not trivial [Ida, 1973]. Smoother solutions can alternatively be achieved by viscous regularization of the frictional response. In particular, rate-and-state friction laws motivated by slow speed experiments [Dieterich; Ruina] operate such regularization [Favreau et al, 1999; Nakatani, 2001; Kaneko et al, 2008].

Earthquake rupture simulations are often triggered by artificial initial conditions that are non-smooth in space and time, such as prescribed abrupt stress drops with piecewise constant spatial distribution [3]. Initiation procedures that are mechanically consistent with nucleation models imply additional complexity and computational cost [Uenishi and Rice, 2003; Ripperger et al., 2008]. Although generally overlooked in earlier work, in many relevant situations nucleation can have a strong impact on large scale dynamic rupture [Nielsen et al; Festa and Vilotte, 2006; Rubin and Ampuero, 2007].

Other irregularities can pose specific difficulties: rupture arrest on abrupt barriers [Voisin et al], interaction of dipping faults with free surface [Madariaga], kinked faults [Madariaga et al. 2006].

1.3   Overview of the numerical method and remarks on the usage of a high order method

We will discuss the implementation of fault dynamics in the Spectral Element Method (SEM). The starting point of our work is the SEM described by Komatitsch and Vilotte [1998], a high-order accurate and flexible method originally introduced in computational fluid dynamics [Patera, 1984] that has been succesfully applied in seismic wave propagation [Komatitsch et al 2005; Chaljub et al 2007]. The contact and friction conditions on the fault interface are solved with the "traction at split nodes" (TSN) method [e.g. Andrews 1975, 1999; Day et al. 2005]. This is the most natural formulation in methods of the finite element family. Most of the presentation in the next section holds for any finite element method with mass lumping.

High-order methods, such as the SEM, feature spectral convergence: discretization errors behave asymptotically as $O(h^{p+1})$, where p is the degree of the polynomial expansion [see for instance Cohen's book for L^2 error estimates; Ampuero and Nissen-Meyer, 2008]. This is often coined as exponential convergence, a term that holds for p-refinement (increasing p at fixed h), simply because $h^p = \exp(p \ln(h))$. This property holds only for infinitely smooth solutions. In more general cases, the error behaves as $O(h^{\min(p,\alpha)})$ where α is the degree of singularity of the physical solution, i.e. the order of its lowest discontinuous derivative. For instance, a jump in slip rate, typically generated by an abrupt stress drop initialization, has $\alpha = 0$ hence the error in slip rate is $O(1)$, i.e. it does not converge (uniformly) upon mesh refinement ! In the presence of sharp gradients and discontinuities the accuracy of spectral methods deteriorates, due to the Gibbs phenomenon: the pointwise convergence of global approximations of discontinuous functions is at most first order [e.g. Deville-Fischer-Mund, or Boyd books].[4]

It is not evident a priori that there are distinct advantages in applying a high-order method to solve non-smooth problems such as those usually encountered in earthquake dynamics modeling. We note however that the spectral formulation of the boundary integral equation

---

[3]cite here some old references and the SCEC tests

[4]nettoyer ce paragraphe, rester bref mais exact

method (Geubelle and Rice; Lapusta et al.; etc), which assumes smooth slip in the evaluation of the stress transfer functionals, has been applied with great success to rupture problems and is the current benchmark method for simple geometries. High-order methods are in a privileged position to achieve low numerical dispersion, which is essential to accurately represent a key ingredient of rupture problems, the elastodynamic stress transfer along a fault. [5]

2 The method

2.1 Statement of the problem

Let's consider a bounded region Ω of the Earth's crust, in which linear elasticity is assumed. The governing linear elastodynamic equations are:

$$\rho\,\mathbf{a} = \mathrm{div}(\boldsymbol{\sigma}) \tag{1}$$
$$\boldsymbol{\sigma} = \mathbf{c} : \boldsymbol{\epsilon} \tag{2}$$
$$\boldsymbol{\epsilon} = (\nabla\mathbf{d} + \nabla^T\mathbf{d})/2. \tag{3}$$

where ρ is density, $\mathbf{d}(\mathbf{x}, t)$ and $\mathbf{a}(\mathbf{x}, t)$ are particle displacements and accelerations respectively, $\boldsymbol{\sigma}$, $\mathbf{c}$ and $\boldsymbol{\epsilon}$ are the stress, elastic moduli and strain tensors respectively, and div is the divergence operator. Small displacements and the Eulerian view are assumed throughout. Null initial conditions are prescribed for $\mathbf{d}$ and $\mathbf{v}$, assuming the medium is initially at static equilibrium. All variables are understood as perturbations relative to the equilibrium fields. Free stress boundary conditions are imposed on the upper surface of Ω and radiation conditions (PML or ABC) are assumed on its remaining external boundaries. For the sake of conciseness we will ignore these external boundary conditions in the remainder of this presentation, which is focused on the fault boundary conditions.

We define a fault surface as an internal interface Γ across which displacement discontinuities are allowed. For the sake of definiteness of the contact and friction conditions, we further assume that the fault is made of a single smooth surface with continuous normal vector $\mathbf{n}$. The appropriate treatment of fault kinks and branching points is formulated elsewhere [6]. The interface Γ is actually made of two surfaces, Γ_+ and Γ_-, in frictional contact: to each geometric point $\mathbf{x}$ located on Γ correspond two material points $\mathbf{x}_\pm$ located on each side of the fault. We define the orientation of $\mathbf{n}$ as going from Γ_+ to Γ_- and the displacements on each material side of the fault as

$$\mathbf{d}_\pm = \mathbf{d}(\mathbf{x}_\pm) \tag{4}$$

The displacement discontinuity across Γ is denoted

$$\Delta\mathbf{d} = \mathbf{d}_+ - \mathbf{d}_- \tag{5}$$

The normal and tangential components of $\Delta\mathbf{d}$ are respectively defined as

$$\Delta d_N = \Delta\mathbf{d} \cdot \mathbf{n} \tag{6}$$
$$\Delta\mathbf{d}_T = \Delta\mathbf{d} - \Delta d_N \mathbf{n} \tag{7}$$

---

[5]is this the main reason of success of SEM ?
[6]add reference, future work

With these conventions Δd_N is negative for separation and positive for penetration between the two fault sides. Similar definitions and notations are adopted for velocities $\mathbf{v}$ and accelerations $\mathbf{a}$. The tractions acting on each side of the fault are denoted

$$\boldsymbol{\tau}_\pm = \pm\boldsymbol{\sigma}_\pm \cdot \mathbf{n} \tag{8}$$

Owing to the action-reaction principle (Newton's third law), in the remainder we will write

$$\boldsymbol{\tau} \doteq \boldsymbol{\tau}_+ = -\boldsymbol{\tau}_-. \tag{9}$$

The normal and tangential components, τ_N and $\boldsymbol{\tau}_T$, are defined as previously, τ_N being negative for compression and positive for tension.

Mixed boundary conditions are prescribed on the fault: friction and contact laws relate $\boldsymbol{\tau}$ to $\Delta\mathbf{d}$, $\Delta\mathbf{v}$ or to other variables describing the fault state. In most CED applications the relation is a non regular graph. This is the case for Coulomb and linear slip-weakening friction and for the Signorini conditions of unilateral contact, described below. Alternative smooth formulations include rate-and-state friction and regularized formulations by viscous or elastic penalty terms, not treated here. The friction and contact conditions apply to the absolute (total) fields, which include the initial values for fault stress, slip and state variables.

The normal components of $\boldsymbol{\tau}$ and $\Delta\mathbf{d}$ are typically related by unilateral contact conditions, also known as Signorini conditions. These imply no interpenetration of the two fault sides and vanishing normal stress on open faults. These conditions can be formulated as

$$\Delta d_N \leq 0 \tag{10}$$
$$\tau_N \leq 0 \tag{11}$$
$$\Delta d_N \, \tau_N = 0 \tag{12}$$

The tangential components are related by friction conditions which imply that shear tractions are bounded by frictional strength proportional to normal tractions, that slip is allowed only when the strength is reached and that slip rate and shear tractions are anti-parallel. These conditions can be formulated as

$$\mu\tau_N + |\boldsymbol{\tau}_T| \leq 0 \tag{13}$$
$$|\Delta\mathbf{v}_T| \, (\mu\tau_N + |\boldsymbol{\tau}_T|) = 0 \tag{14}$$
$$\Delta\mathbf{v}_T \, |\boldsymbol{\tau}_T| + |\Delta\mathbf{v}_T| \, \boldsymbol{\tau}_T = 0 \tag{15}$$

where the friction coefficient $\mu \geq 0$ can depend on $|\Delta\mathbf{d}_T|$, $|\Delta\mathbf{v}_T|$ or on other fault state variables. A basic example is the linear slip-weakening friction law without healing:

$$\mu = \min\left[\mu_s - (\mu_s - \mu_d)\frac{\delta}{D_c}, \mu_d\right] \tag{16}$$
$$\dot{\delta} = |\Delta\mathbf{v}_T| \tag{17}$$

The fault boundary conditions should in principle be enforced along the deformed interface, accounting in particular for rotations of the normal vector $\mathbf{n}$. However, this complication has not been considered in usual pratice. As noted by Caroli and Velicky [2003], the resulting correction terms are small, of same order as the ratio of initial stresses to elastic moduli ($\approx 10^{-3}$ in CED applications), and can be safely neglected.

2.2   Variational formulation

The SEM, like the FEM, is based on a variational formulation of the problem. The classical formulation as a variational inequality (e.g. Duvaut and Lions, 1972) is not the most suitable for implementation. Here we work with a hybrid variational formulation, with Lagrange multipliers, in which displacements and fault tractions are primary unknowns. The contact and friction conditions are enforced in a weak sense. [7]

We define the following functional spaces:

$$
\begin{aligned}
\mathcal{S} &= H^1(\Omega)^3 \\
\mathcal{S}_\tau &= \{\boldsymbol{\tau} \in H^{1/2}(\Gamma)^2 \mid \forall \boldsymbol{\lambda} \in H^{-1/2}(\Gamma)^2, \\
&\quad (\tau_N, \lambda_N) \le 0 \text{ and} \\
&\quad |(\boldsymbol{\tau}_T, \boldsymbol{\lambda}_T)| + (\mu\tau_N, \lambda_N) \le 0\}
\end{aligned}
$$

(18)

(19)

where $H^\alpha(X)$ is the space of square-integrable scalar functions of $\mathbf{x} \in X$ with square-integrable generalized derivatives of order α in the sense of distributions and $H^{-\alpha}(X)$ is its dual space (see e.g. Duvaut and Lions, 1972). We introduce the bilinear form $a(\cdot, \cdot)$ in $\mathcal{S}$ expressing the virtual work of internal stresses:

$$
a(\mathbf{u}, \mathbf{w}) \doteq \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \nabla\mathbf{w} \, dV = \int_\Omega \nabla\mathbf{u} : \mathbf{c} : \nabla\mathbf{w} \, dV
$$

(20)

and denote the dot product on $\mathcal{S}$ by

$$
(\mathbf{u}, \mathbf{w}) = \int_\Omega \mathbf{u} \cdot \mathbf{w} \, dV
$$

(21)

and the dot product on Γ by

$$
(\boldsymbol{\lambda}, \boldsymbol{\mu})_\Gamma = \int_\Gamma \boldsymbol{\lambda} \cdot \boldsymbol{\mu} \, dS.
$$

(22)

The variational formulation of our problem within the time interval $[0, T]$ is:

Find $\mathbf{d}(\mathbf{x}, t)$ in $\mathcal{S} \times [0, T]$ and $\boldsymbol{\tau}(\mathbf{x}, t)$ in $\mathcal{S}_\tau \times [0, T]$ such that for all $\mathbf{w} \in \mathcal{S}$ and $\boldsymbol{\lambda} \in \mathcal{S}_\tau$ we have

$$
\begin{aligned}
(\rho\ddot{\mathbf{d}}, \mathbf{w}) &= -a(\mathbf{d}, \mathbf{w}) + (\boldsymbol{\tau}, \Delta\mathbf{w})_\Gamma \\
(\boldsymbol{\tau} - \boldsymbol{\lambda}, \Delta\mathbf{d})_\Gamma &\le 0
\end{aligned}
$$

(23)

(24)

with initial conditions

$$
\begin{aligned}
(\mathbf{d}(\mathbf{x}, 0), \mathbf{w}) &= (\mathbf{d}_0, \mathbf{w}) \\
(\mathbf{v}(\mathbf{x}, 0), \mathbf{w}) &= (\mathbf{v}_0, \mathbf{w})
\end{aligned}
$$

(25)

(26)

---

[7] C'est ce qu'on prétend. Dans la suite j'ai essayé de faire une présentation formelle mais pas trop lourde, mais certains passages restent obscurs. J'ai principalement deux doutes: quel est l'espace approprié pour les tractions (éviter un "crime variationnel")? est-il nécessaire d'être très rigoureux sur la formulation faible des conditions de contact et frottement alors que l'implémentation finale semble équivalente à la formulation forte?

The existence and uniqueness of solutions for the general Signorini problem with friction are still open questions, although results are available for some special cases (small friction coefficient; for an overview see Renard, 2005, Coorevits et al. 2001) and for related problems (Tresca problem, in which fault strength is independent of normal stress). At this point we adopt a pragmatic approach ... [8]

2.3   Spatial discretization and algebraic formulation

The domain Ω is decomposed into n_e non-overlapping hexaedral elements Ω_e. We define a mesh size parameter h as the largest element size (radius of containing sphere). The mesh can be unstructured and the elements deformed as required to honor the prominent interfaces of the geological model, including the fault geometry (figure XXX). The faults are located along element boundaries and have split nodes to accomodate displacement discontinuities: each geographical point on a fault surface is made of two material nodes facing each other. Note that we adopt a Eulerian description of motion so the computational mesh is fixed, even on the fault nodes.

We denote by $\mathcal{F}_e$ the invertible coordinate mapping from the reference cube $[-1, 1]^3$ to the element Ω_e such that $\mathbf{x} = \mathcal{F}_e(\boldsymbol{\xi})$, $L^2(\Omega)$ is the space of square-integrable scalar functions of $\mathbf{x} \in \Omega$ and P_p the space of polynomials of degree up to p on the reference segment $[-1, 1]$. The functional space of displacements $\mathcal{S}$ is approximated by

$$\mathcal{S}^{hp} = \{\mathbf{u} \in L^2(\Omega)^3 \mid \forall e \ \mathbf{u}|_{\Omega_e} \circ \mathcal{F}_e \in P_p^3\} \tag{27}$$

The D-dimensional polynomial basis adopted in the reference element $[-1, 1]^D$ is the tensor product of the Lagrange polynomials through the Gauss-Lobatto-Legendre (GLL) points of degree p on $[-1, 1]$. This implies that each element is endowed with an internal regular mesh of $(p+1)^D$ GLL nodes. The integrals in the variational formulation are approximated by the tensorized GLL quadrature.

The space of tractions $\mathcal{S}_\tau$ is approximated by[9]

$$\mathcal{S}_\tau^{hp} = \{\boldsymbol{\tau} \in L^2(\Gamma)^3 \mid \forall e \ \boldsymbol{\tau}|_{\Gamma_e} \circ \mathcal{F}_e \in P_p^2\} \tag{28}$$

In the remainder $\mathbf{d}$ denotes the vector collecting the nodal values of displacements of all the computational nodes of the bulk mesh, $\Delta\mathbf{d}$ the vector collecting the difference between values of corresponding split nodes on Γ_+ and $\Gamma-$, and $\boldsymbol{\tau}$ the vector collecting the tractions on the nodes of the fault surface Γ_+. Note that these vectors include values at the internal GLL nodes of each element and are non-redundant at the common nodes on element boundaries. The semi-discrete problem is (e.g. Coorevits et al ., 2001):

Find for $t \in [0, T]$ the displacements $\mathbf{d}(t)$ and tractions $\boldsymbol{\tau}(t)$ such that

$$\mathbf{Ma} = -\mathbf{Kd} + \mathbf{B}\boldsymbol{\tau} \tag{29}$$

---

[8]Discussion à compléter.

[9]Dans l'implémentation on a utilisé le même ordre polynomial pour les tractions et pour le volume. Est-ce une violation de la condition de Ladysenskaja-Babuska-Brezzi ?

and $\boldsymbol{\tau}$, $\Delta\mathbf{d}$ and $\Delta\mathbf{v}$ satisfy the relations (10)-(15) on a node-by-node basis. [10]

Here $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices, given for instance by Komatitsch and Vilotte (1998), and $\mathbf{B}$ is a fault boundary matrix given explicitly in Appendix A. Note that usual implementations with direct stiffness summation do not require the storage of $\mathbf{K}$. Due to the choice of interpolation nodes coinciding with the quadrature nodes $\mathbf{M}$ and $\mathbf{B}$ are diagonal, a property that is essential for the efficiency of the SEM.

2.4 Time discretization and implementation details

Wave propagation problems in SEM are usually discretized by an explicit Newmark algorithm or an explicit Hilber-Hughes-Taylor (HHT-α) algorithm. To simplify the presentation we take here the particular case of a fully explicit mid-point rule scheme ($\alpha = 1$, $\beta = 0$ and $\gamma = 1/2$, this is the scheme in the SPECFEM3D code):

$$\mathbf{d}_{n+1} = \mathbf{d}_n + \Delta t \mathbf{v}_n + \frac{\Delta t^2}{2}\mathbf{a}_n \tag{30}$$

$$\mathbf{M}\mathbf{a}_{n+1} = -\mathbf{K}\mathbf{d}_{n+1} + \mathbf{B}\boldsymbol{\tau}_{n+1} \tag{31}$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \Delta t \frac{\mathbf{a}_n + \mathbf{a}_{n+1}}{2} \tag{32}$$

Similar algorithms can be derived for other time discretization schemes by following the steps outlined in the remainder of this section.

The procedure to advance one timestep is summarized in Figure 1. The critical stage is contained in steps 3-a and 3-b, where the fault tractions $\boldsymbol{\tau}_{n+1}$ are obtained by solving the elastodynamic equation 31 on the fault simultaneously with the fault boundary conditions. Combining equations 31 and 32:

$$\mathbf{v}_{n+1} = \tilde{\mathbf{v}}_{n+1} + \frac{\Delta t}{2}\mathbf{M}^{-1}\mathbf{B}\boldsymbol{\tau}_{n+1} \tag{33}$$

where

$$\tilde{\mathbf{v}}_{n+1} = \mathbf{v}_n + \frac{\Delta t}{2}(\mathbf{a}_n - \mathbf{M}^{-1}\mathbf{K}\mathbf{d}_{n+1}) \tag{34}$$

is the "free velocity" that would prevail if the fault tractions were suddenly cancelled to get a free boundary condition. We now specialize equation 33 to the fault nodes:

$$\mathbf{v}_{n+1}^{\pm} = \tilde{\mathbf{v}}_{n+1}^{\pm} \pm \frac{\Delta t}{2}\mathbf{M}_{\pm}^{-1}\mathbf{B}_{\pm}\boldsymbol{\tau}_{n+1} \tag{35}$$

Subtracting both sides ($\pm$) leads to

$$\Delta\mathbf{v}_{n+1} = \Delta\tilde{\mathbf{v}}_{n+1} + \mathbf{Z}^{-1}\boldsymbol{\tau}_{n+1} \tag{36}$$

where $\mathbf{Z}$ is a fault impedance matrix defined by

$$\mathbf{Z}^{-1} = \frac{\Delta t}{2}(\mathbf{M}_{+}^{-1}\mathbf{B}_{+} + \mathbf{M}_{-}^{-1}\mathbf{B}_{-}) \tag{37}$$

---

[10]La formulation de départ des conditions de contact et frottement se veut variationnelle, mais ici je me retrouve avec une condition locale forte (pointwise) sur la faille. Certains références suggèrent que ceci n'est pas optimal (ou "variational crime"?). Peut-on dire que les formulations faible et forte des conditions d'interface sont équivalentes grâce à la diagonalité des matrices $\mathbf{M}$ et $\mathbf{B}$?

1. Displacement update and partial velocity update:

$$\mathbf{d} \leftarrow \mathbf{d} + \Delta t \mathbf{v} + \frac{\Delta t^2}{2} \mathbf{a}$$

$$\mathbf{v} \leftarrow \mathbf{v} + \frac{\Delta t}{2} \mathbf{a}$$

2. Compute the internal forces:

$$\mathbf{f} = -\mathbf{K} \mathbf{d}$$

3. Apply the dynamic fault boundary conditions:

   (a) Compute the fault stick tractions

   $$\tilde{\boldsymbol{\tau}} = -\mathbf{Z} \, \Delta \tilde{\mathbf{v}}$$

   with $\tilde{\mathbf{v}} = \mathbf{v} + \frac{\Delta t}{2} \mathbf{M}^{-1} \mathbf{f}$ (only on fault nodes)

   (b) Find $\boldsymbol{\tau}$ by solving

   $$\boldsymbol{\tau} = \tilde{\boldsymbol{\tau}} + \mathbf{Z} \, \Delta \mathbf{v}$$

   together with the contact and friction conditions (10)-(15)

   (c) Add the fault boundary term to the sum of forces:

   $$\mathbf{f} \leftarrow \mathbf{f} + \mathbf{B} \boldsymbol{\tau}$$

4. Solve for acceleration:

$$\mathbf{a} \leftarrow \mathbf{M}^{-1} \mathbf{f}$$

5. Final velocity update:

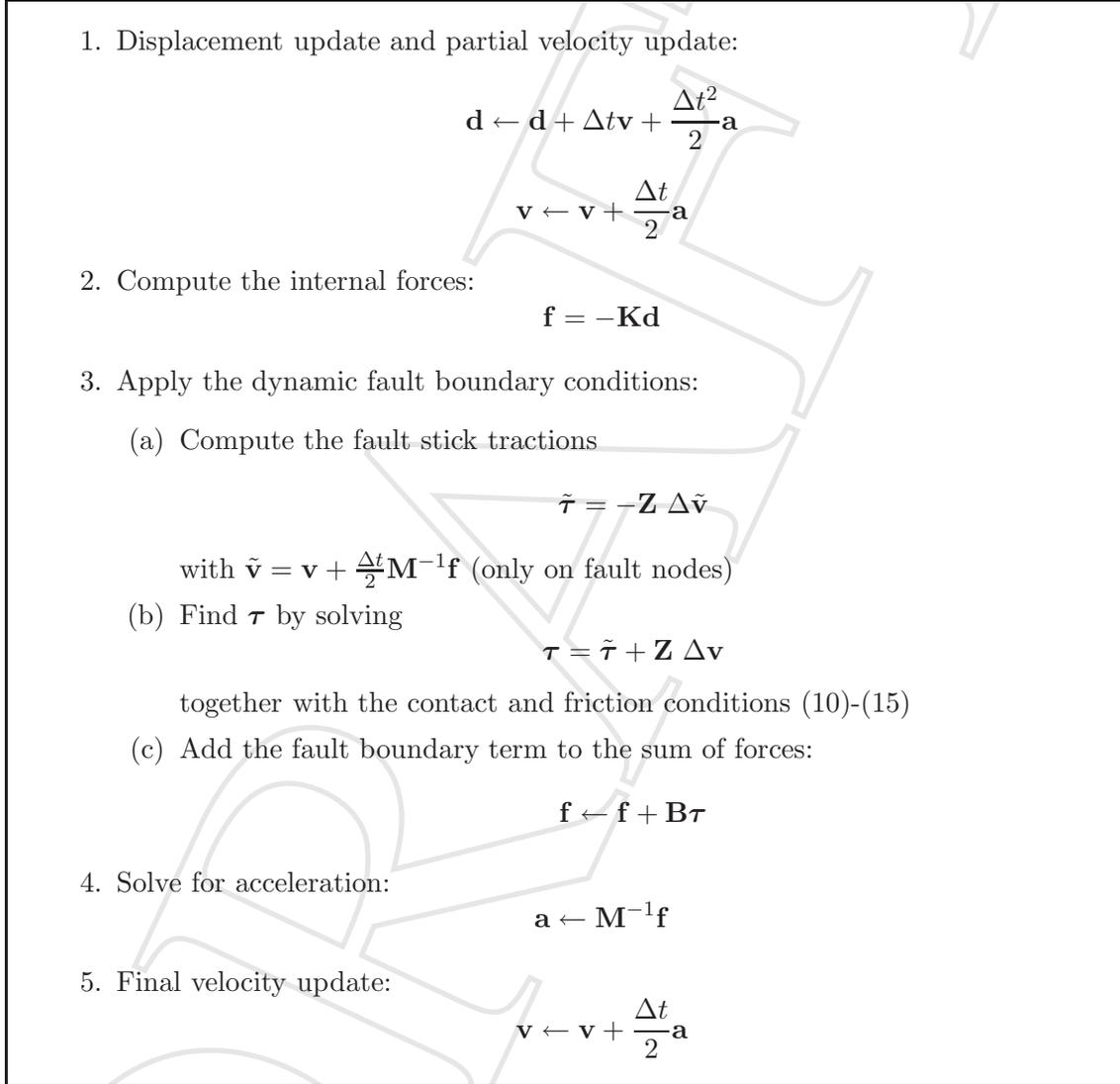$$\mathbf{v} \leftarrow \mathbf{v} + \frac{\Delta t}{2} \mathbf{a}$$

Figure 1: Timestepping algorithm with fault frictional contact solver (step 3) embedded in an explicit Newmark scheme. Step 4 depends on the assumed fault constitutive laws.

Note that for conformal meshes $\mathbf{B}_+ = \mathbf{B}_-$. It is convenient to rewrite 36 as

$$\boldsymbol{\tau}_{n+1} = \tilde{\boldsymbol{\tau}}_{n+1} + \mathbf{Z}\,\Delta\mathbf{v}_{n+1} \tag{38}$$

where

$$\tilde{\boldsymbol{\tau}}_{n+1} = -\mathbf{Z}\,\Delta\tilde{\mathbf{v}}_{n+1} \tag{39}$$

is the "stick traction" that would immediately prevail if slip arrest was suddenly enforced ($\Delta\mathbf{v}_{n+1} = 0$). Note that in (34) and (39) the update of $\tilde{\boldsymbol{\tau}}_{n+1}$ is fully explicit. Equation 38 is the affine relation between slip rate and traction arising from the SEM discretization of the bulk equations. This relation must be solved simultaneously with the fault constitutive laws (10)-(15) to obtain $\boldsymbol{\tau}_{n+1}$, as will be detailed next. These steps are treated efficiently in SEM because the fault impedance matrix $\mathbf{Z}$ is diagonal: equation 38 is a local relation that can be solved node by node on the fault.

We drop the $n+1$ index and note by $\mathcal{T} = \tau + \tau^0$ the total tractions. Solving the unilateral contact conditions (10)-(12) on each fault node amounts to find the intersection of the graph in figure XXX with

$$\mathcal{T}^N = \tilde{\mathcal{T}}^N + Z\,\Delta v^N \tag{40}$$

The solution is (see figure XXX)

$$\mathcal{T}^N = \min(0, \tilde{\mathcal{T}}^N)\,H(\Delta d^N) \tag{41}$$

where H is the Heaviside function ($H(x) = 0$ if $x < 0$ else $H(x) = 1$). [11]

If the friction coefficient μ is known at this stage, the solution of

$$\mathcal{T}^T = \tilde{\mathcal{T}}^T + Z\,\Delta v^T \tag{42}$$

with the friction conditions (13)-(15) is:

$$\mathcal{T}^T = \min(-\mu\mathcal{T}^N, \|\tilde{\mathcal{T}}^T\|)\,\frac{\tilde{\mathcal{T}}^T}{\|\tilde{\mathcal{T}}^T\|} \tag{43}$$

3   Characteristics of high frequency numerical noise

3.1   Numerical examples

We solve a 2D mode III version of the SCEC Problem Version 3. We compare the SEM to a spectral implementation of the boundary integral equation method that follows closely Lapusta et al (2000). [...] Emphasize the appearance of high frequency numerical noise and its effect on the convergence rate (i.e. on the exponent of the reduction of rupture time error as a function of decreasing mesh size h).

The remainder of this section is devoted to examine the origin of the high frequency noise and to propose and evaluate different methods to control it.

The spurious noise is better characterized on a 1D nucleation test. Consider a planar fault in an unbounded homogeneous elastic medium, governed by linear slip-weakening friction. Frictional properties and initial stresses are uniform along the fault. The initial shear stress

---

[11]Warning: due to the fully explicit update of displacements this time scheme does not guarantee the non-penetration condition at the next timestep !

is slightly above the static threshold, $\tau_0 = \tau_s + \epsilon(\tau_s - \tau_d)$ with $0 < \epsilon \ll 1$. The resulting motion is spatially uniform along the fault plane directions and the problem is one-dimensional. It can be solved analytically by equating the frictional strength to the fault stress, which in 1D reduces to a radiation damping response:

$$\tau = \tau_0 - \frac{\mu}{2c_S}V = \begin{cases} \tau_s - (\tau_s - \tau_d)D/D_c & \text{if } D < D_c \\ \tau_d & \text{otherwise} \end{cases} \tag{44}$$

The resulting slip velocity has an initial jump, followed by an exponential growth, and finally reaches a constant value:

$$\frac{V(t)}{s_m D_c} = \begin{cases} 0 & \text{if } t < 0 \\ \epsilon \exp(s_m t) & \text{if } 0 \leq t < T_\nu \\ 1 + \epsilon & \text{otherwise} \end{cases} \tag{45}$$

where

$$s_m = \frac{2c_S(\tau_s - \tau_d)}{\mu D_c} \tag{46}$$

$$T_\nu = s_m^{-1} \ln(1 + 1/\epsilon) \tag{47}$$

This solution contains two irregularities of different order: a discontinuity of slip velocity at $t = 0$ due to the abrupt initial conditions, and a discontinuity of slip acceleration at $t = T_\nu$ due to the discontinuous slope of the slip-weakening law. The particle velocity at a location z outside the fault is

$$v(z, t) = sign(z) \, V(t - |z|/c_S)/2 \tag{48}$$

A characteristic length-scale of the velocity distribution in the off-fault direction is

$$\lambda_\perp = c_S/s_m = \frac{\mu D_c}{2(\tau_s - \tau_d)} \tag{49}$$

Solutions (45) and (48) are also valid if the nucleation region, the fault or the domain are bounded, as long as no wave reflected or diffracted by the boundaries has arrived to the selected location.

Figure XXX compares the analytical slip velocity (Equation 45 with $\epsilon = 0.1$) to the numerical solution obtained with a 2D SEM code for SH waves (boundaries perpendicular and parallel to the fault are periodic and absorbing, respectively). The element size was set to $h = \lambda_\perp$ and the polynomial degree to $p = 5$. The numerical solution follows the correct trend, but superimposed are spurious oscillations. A computation including artificial numerical damping (Figure XXX) shows that these oscillations are generated by the two discontinuities of the solution. Their frequency is grid-dependent, proportional to element size h. The grid-dependency of their amplitude depends on the order of the discontinuity: the oscillation amplitudes generated by the stronger discontinuity ($t = 0$) remain constant when h is reduced, whereas the oscillations induced by the weaker discontinuity ($t = T_\nu$) have amplitudes proportional to h. Reducing the timestep at fixed h does not affect these properties, and increasing p has a similar effect than reducing h. The spatial pattern of the spurious oscillation modes is trapped near the fault (Figure XXX).
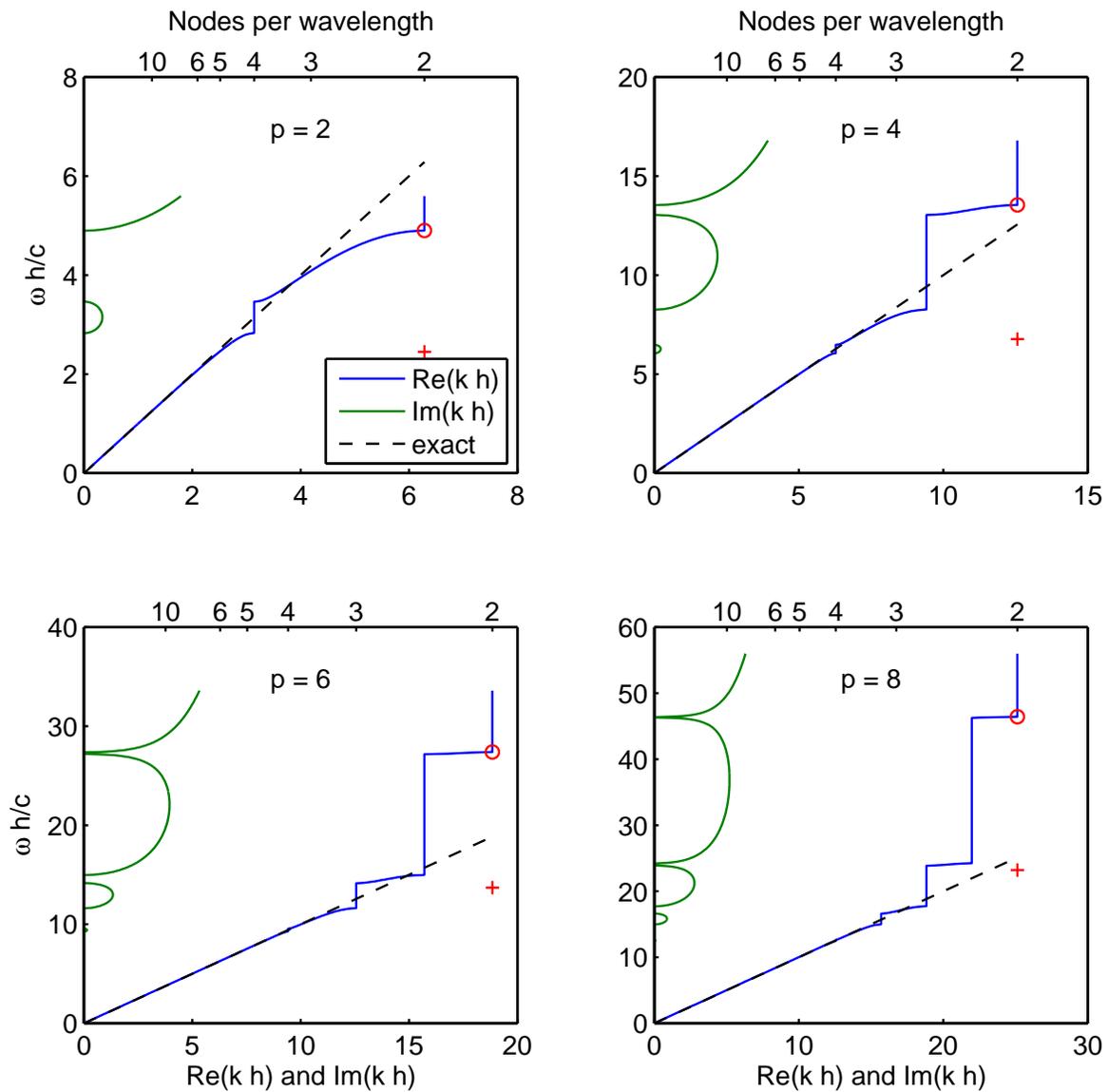
Figure 2: Dispersion curves (1D) for SEM, including complex wavenumbers. The red circle indicates the highest propagative mode, which gives the stability limit in explicit time schemes. The red cross indicates half that frequency.

3.2   Theoretical analysis of the method: numerical dispersion and impedance

The numerical dispersion properties of the SEM were recently summarized by [Ampuero and Nissen-Meyer, 2008]. Appendix B outlines the derivation of the discrete 1D dispersion curves, the relation between frequency ω and wavenumber k (possibly complex). Figure 2 shows dispersion curves for four usual values of the polynomial order p. Within specific frequency bands, called *stopping bands* [Thompson and Pinsky, 1994], the numerical waves are inhomogeneous: their wavenumbers have an imaginary part that produces an exponential decay in space. There are p stopping bands, located at $kh = j\pi$ with $j = 1$ to p, of which $p - 1$ are within the stable wavenumber range ($kh < p\pi$). These modes have an elegant interpretation in the context of wave propagation in discrete lattices, as exposed by Brillouin (1953).

The discrete 1D boundary impedance $Z(\omega)$, defined as the spectral ratio between traction and velocity at the boundary of a 1D half-space, is derived in Appendix B and plotted as a function of frequency in Figure 3. In the low frequency range the discrete impedance is close to the analytical one (the error behaves as $(kh)^{2p}$). At higher frequencies the impedance has troughs and peaks occuring at the characteristic frequencies that define the edges of the stopping bands of the dispersion relation, as shown in Appendix B. Under time-dependent Dirichlet (resp. Neumann) boundary conditions the peaks (resp. troughs) are resonant frequencies that produce self-sustained oscillations of the boundary tractions (resp. displacements). The amplitude of these oscillations is determined by the spectral power of the source time function at the characteristic frequencies, and is naturally higher for broad band sources. The characteristic modes decay exponentially in the normal direction to the boundary. The first few stopping bands (low frequency) are very narrow and have very slow spatial decay ($Im(kh) \ll \pi$), so the associated resonant modes are not a nuisance in practice. However, the higher stopping bands, corresponding to less than 3 to 4 nodes per wavelength (Figure 2), are particularly problematic: these are at the origin of spurious inhomogeneous modes trapped in the vicinity of the fault.

4   Reduction of high frequency numerical noise

4.1   Dissipative time integration schemes

The following estimate indicates that the spurious frequencies are typically much larger than the sampling frequency $1/\Delta t$. For $p = 4$ the first significant spurious frequency lies close to one half of the cut-off frequency for S wave propagation, ω_{max}^S (the edge of the last stopping band). Stability of explicit time integration requires $\omega_{max}^P \Delta t \leq C/\sqrt{D}$, where C is the CFL number ($C = 2$ for a second-order centered scheme) and D is the dimensionality of the problem. Considering $\omega_{max}^P/\omega_{max}^S = c_P/c_S$ is the ratio of P to S wave velocities, the characteristic period of this spurious mode is larger than $4\pi\sqrt{D}/C\ c_P/c_S\ \Delta t$. For usual values of Poisson ratio this is an order of magnitude larger than Δt, and can be much larger within damaged fault zones with high c_P/c_S.

Explicit time schemes can be especially designed to include dissipation at high frequencies, at the expense of introducing distortions at low frequencies. Hulbert and Chung [1996] proposed the explicit generalized-α method, an extension of the Newmark-α (or HHT-α) family optimized in order to maximize high frequency dissipation (controlled by the parameter α) while minimizing low frequency dispersion. The friction and contact solver can be readily
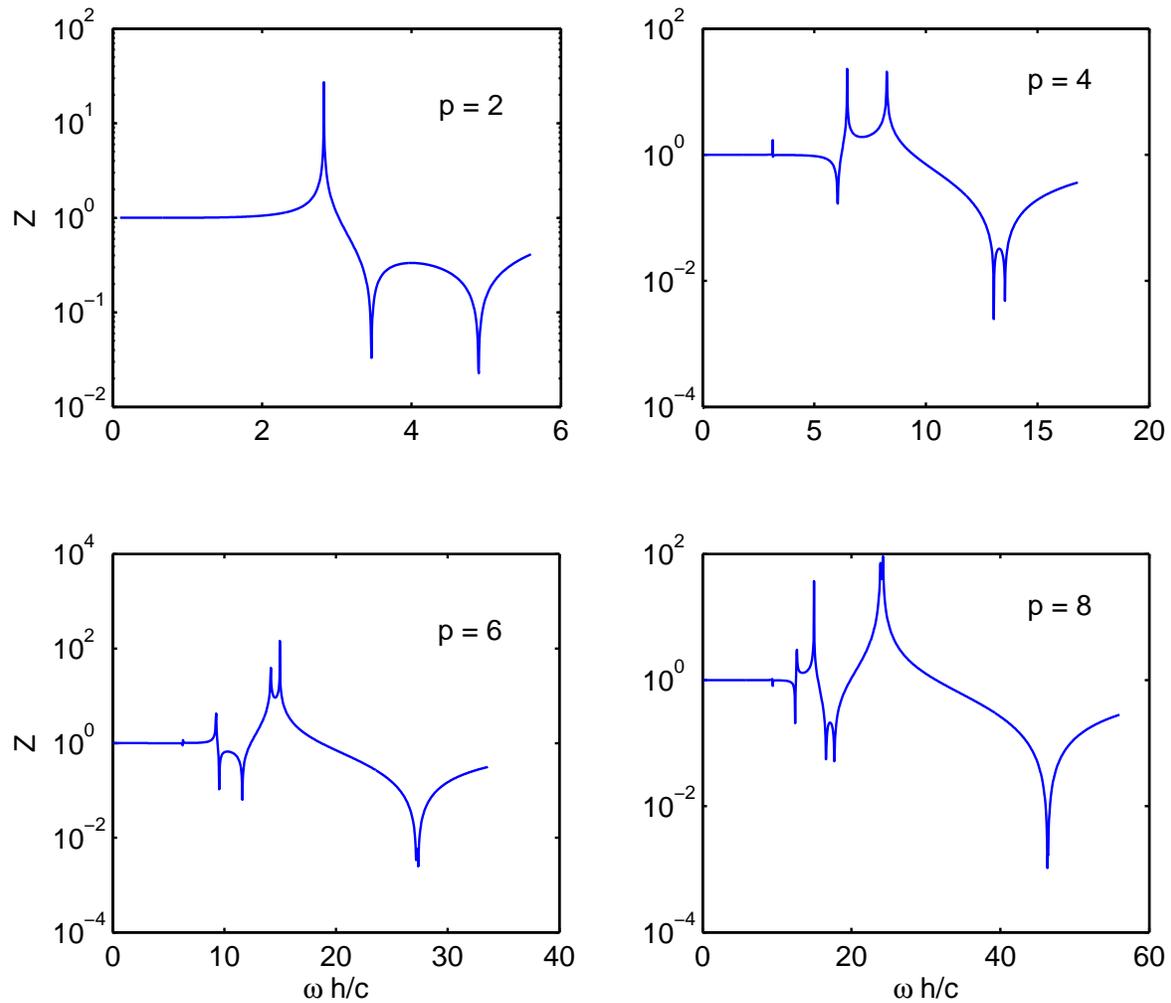
Figure 3: Amplitude of the discrete boundary impedance (1D) for SEM as a function of frequency ω, for four usual polynomial degrees. The analytical value is $Z = 1$.

embedded in this timestepping algorithm following the same lines described in the previous section. However, its efficiency in controlling the spurious modes is limited. Significant numerical damping can be only introduced in the frequency range close to the critical sampling frequency ω_{max}^P, which determines the timestep Δt, but the spurious frequencies are generally much lower and are damped too slowly. This defficiency is exhacerbated in complex geometries that require meshes with elements of a broad range of sizes close to the fault, as Δt and the dissipative frequency band are proportional to the smallest element size whereas the lowest spurious frequency is proportional to the largest element size.

The SEM could be applied with unconditionally stable implicit timestepping schemes, with optimal high frequency dissipation, for which it is possible to set $1/\Delta t$ below the stopping bands to guarantee damping of the spurious modes. Again, especially designed algorithms with steep dissipation spectrum must be selected to minimize distortion in the usable low frequency band. Implicit methods involve resolution of large matrix systems at each time step and the trade-off between computational cost and accuracy must be carefully assessed. Moreover, a heavier implementation complexity is implied by the need for good preconditioners and by the implicit treatment of the contact and friction conditions. Implicit timestepping has been previously used in SEM for wave propagation [Seriani; Zampieri and Pavarino]. Very recently, Kaneko, Lapusta and Ampuero [AGU 2007] implemented an implicit scheme in SEM for rupture problems that involve multiple timescales, from slow (interseismic) to fast (coseismic). However, an optimal implementation with additive Schwarz preconditioning, not yet available, is needed to assess the relative advantages compared to explicit methods on problems involving only dynamic timescales.

4.2   Spectral filtering

Spatial filtering is often applied to stabilize CFD simulations. Selective modal damping is also applied to control hourglass modes in FEM [Daniel and Belytschko, 2005]. Advantage, in principle: it can be tuned to each element size. Describe and report main results. [...]

To ensure an element-by-element implementation, filtering must be applied to bubble modes with zero displacement at the edges of the elements (homogeneous Dirichlet modes). This type of filtering cannot deal at the same time with peaks and troughs of the discrete impedance, because these are associated to the resonance of distinct Dirichlet and Neumann modes, respectively.

4.3   Kelvin-Voigt damping

Damping by an artificial viscous material has been proposed as an adequate solution for the control of high frequency artefacts by Knopoff and Ni [2001] and Day and Ely [2002].

Kelvin-Voigt viscous material [Day and Ely, 2002]: a viscous term $\mathbf{Cv}$ is added to the discrete elastodynamic equations, with $\mathbf{C} = \eta\mathbf{K}$, to introduce a frequency dependent quality factor $Q^{-1} = \eta\,\omega$. The viscosity parameter η is fixed such that $Q = 1$ at the Nyquist frequency $1/2\Delta t_c$ associated with the critical time-step (for the leap-frog scheme $\Delta t_c = 2/\sqrt{D}\omega_{max}^P$). The normalized viscosity $\eta^* = \eta/\Delta t_c$ is then $\eta^* = 1/\pi \approx 0.318$. At frequencies typical of stopping bands, $\omega \approx 0.1 \times 2\pi/\Delta t_c$, we can expect a decay below 10% of the initial amplitude only after 8 cycles. Knopoff and Ni (2001) proposed an equivalent idea. Their choice is $\eta = \sqrt{0.1}/\omega_c$, which for the leapfrog scheme gives $\eta^* = \sqrt{0.1}/2 \approx 0.158$. [12]

---

[12] Jean Elkhoury (appendix in his Ph.D. dissertation) recently showed that the optimal value of η^* is not

Ben-Jemaa et al (2006, in review) also included a similar damping.

There is some discussion about Kelvin dissipation and rupture dynamics in Marder (2006).

To guarantee stability the critical Δt must be reduced by a factor $1/\sqrt{1 + 2\eta^*}$, which is close to unity for small η^*. Numerical analysis of the time scheme: the eigenvalues of the amplification matrix are

$$\rho_\pm = 1 - (1 + \eta^*)\Omega^2/2 \pm \Omega/2\sqrt{(1 + \eta^*)^2\Omega^2 - 4} \tag{50}$$

where $\Omega = \omega\Delta t$. The stability limit corresponds to $\rho_- = -1$.

To avoid attenuation of meaningful lower frequencies we confine the viscous material to a narrow layer surrounding the fault, with a width of 4 to 6 GLL nodes and η smoothly decaying as a Gaussian with distance normal to the fault plane. [...] Describe main numerical results. Show and discuss the effect on the discrete impedance.

4.4 PML damping

Soon after the introduction of the PML (Berenger 1994), a modified version, the Complex Frequency Shifted PML (Kuzuoglu and Mittra 1996), was developped and shown to absorb evanescent waves propagating parallel to interfaces that appear at the end of electro-magnetic waveguides (Berenger 2002). More recently, Festa et al [2005] and Festa and Vilotte [2006] applied a similar idea, coined GFPML (generalized frequency dependent PML), to the absorption of outgoing surface waves and to the numerical damping of spurious artifacts in fault dynamics, with a convolutional implementation similar to that of Roden and Gedney (2000).

[...] Describe main numerical results. Show and discuss the effect on the discrete impedance.

5 Performance of the method

Convergence in the 3D SCEC benchmark and comparison to other methods. Show convergence results and quantitative comparison to FDM and SBIEM by Day et al. (2005).

Discuss the smoothness of the benchmark case. Show results of a smooth test, the rate-and-state simulations by [Kaneko, Lapusta and Ampuero, 2007, subm. to JGR]. Emphasize that the order of convergence is higher in the smooth case: the advantages of high order methods are fully exploited in physical problems with smooth solution.

Still open question: what is the optimal polynomial order ?

6 Examples and perspectives

This section illustrates through brief examples the potentials of the method.

6.1 Fault surrounded by heterogeneous medium

Show results of rupture on a fault inside a low velocity layer (a damaged fault zone) as in [Harris and Day, 1997]. This configuration highlights the importance of an accurate (low dispersion) modeling of the radiated wavefield, because reflected waves interact with dynamic rupture. Compare SH case to the SBIEM modeling by Hendrickx, Geubelle and Sottos (2005).

---

universal but depends on the rupture speed.

[Kaneko, Lapusta and Ampuero, 2007, subm. to JGR] studied the role of a shallow velocity structure in suppressing supershear rupture speeds near the surface.

6.2 General geometries

Mention published and previously reported applications:

- Supershear and nucleation: Festa and Vilotte (2006).

- Kinked faults: cite Festa and Vilotte (AGU/EGU abstracts), Madariaga et al (2006).

- Rate-and-state: Kaneko et al. (2007, submitted to JGR)

- Faults with linear compliance law: Haney et al. (2007).

Smoothly curved faults (in Festa's Ph.D.). Maybe compare to Victor results.
Maybe do a Landers simulation in 2.5D (in the sense of Johnson [1992]).
Discuss challenges in realistic 3D applications: issues with mesh generation.

6.3 Coupling to general rheologies

Mention our recent work on rupture on faults surrounded by visco-plastic material (as in [Andrews, 2005]) and damage material (as in [Lyakhovsky 2001]).

7 Conclusions

Acknowledgements

A Fault boundary matrix

The fault Γ naturally inherits from the hexaedral element mesh a surface mesh of quadrilateral elements Γ^e, made of all the element faces lying on any of the two sides $\Gamma_\pm$ of the fault. The matrix $\mathbf{B}$ that arises from the discretization of the virtual work of tractions on the fault is obtained by assembling the contributions $\mathbf{B}^e$ from each of these fault boundary elements Γ^e. Note that boundary elements of Γ_+ and Γ_- that face each other are not assembled together. The diagonal term of $\mathbf{B}^e$ associated to any degree of freedom of the GLL node with local indices (i,j) in $\Gamma^e \in \Gamma_\pm$ is

$$B_{ij,ij}^\pm = \pm\, \omega_i \omega_j J_e^{ij} \tag{51}$$

where ω_k are the GLL quadrature weights and

$$J_e^{ij} = \left\| \frac{\partial \mathbf{x}}{\partial \xi} \times \frac{\partial \mathbf{x}}{\partial \eta} \right\| (\xi_i, \eta_j) \tag{52}$$

is the Jacobian of the mapping from the global coordinates $\mathbf{x}$ to the local coordinates $\boldsymbol{\xi} = (\xi, \eta)$ on the reference unit surface element $[-1, 1]^2$.

The normal vector is obtained from the boundary elements of Γ^+ by

$$\mathbf{n} = \frac{1}{J_e} \frac{\partial \mathbf{x}}{\partial \xi} \times \frac{\partial \mathbf{x}}{\partial \eta} \tag{53}$$

The implementation of faults in existing codes is facilitated by the fact that most of the infrastructure needed to compute the quantities above is already in place: they also appear in the implementation of absorbing boundary conditions, as outlined for instance by Komatitsch and Tromp (1999).

B   Fault impedance

We consider a semi-infinite 1D medium ($x \geq 0$) with uniform wave speed c. A Neumann boundary condition, a prescribed traction $\tau(t)$, is applied at $x = 0$. The domain is discretized by a regular mesh of spectral elements with equal size h. We consider the elastodynamic problem in the frequency (ω) domain. The elementary contribution to the dynamic matrix,

$$\mathbf{S} = -\omega^2 \, \mathbf{M} + \mathbf{K} \tag{54}$$

(a symmetric matrix of size $p + 1 \times p + 1$), can be partitioned into the $p - 1$ interior nodes (subscript i) and the 2 boundary nodes of the element (subscript b),

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{bb} & \mathbf{S}_{bi} \\ \mathbf{S}_{ib} & \mathbf{S}_{ii} \end{pmatrix}, \tag{55}$$

then condensed on the boundary nodes (size 2×2):

$$\mathbf{s} = \mathbf{S}_{bb} - \mathbf{S}_{bi} \mathbf{S}_{ii}^{-1} \mathbf{S}_{ib} \tag{56}$$

We denote $d_n(\omega)$ the spectral displacement of the n-th edge node ($x = nh$) and

$$G_1 = s_{12} \tag{57}$$
$$G_2 = -(s_{11} + s_{22})/2. \tag{58}$$

The coefficients G_i are rational functions of $(\omega h/c)^2$. G_1 has $p - 1$ poles and no zero. G_2 shares the same $p - 1$ poles and has p zeros. The condensed discrete problem reads

$$-G_2 d_0 + G_1 d_1 = -\tau \tag{59}$$
$$G_1 d_{n-1} - 2G_2 d_n + G_1 d_{n+1} = 0 \text{ for } n \geq 1 \tag{60}$$

The structure of the condensed problem is similar to that of a finite difference scheme, and the standard Von Neumann analysis technique can be applied. Assuming $d_n(\omega) = d_0(\omega) \exp(ikhn)$, Equation 60 yields the following a relation between wavenumber k and frequency ω, a dispersion relation:

$$\cos(kh) = G_2/G_1 \tag{61}$$

The discrete boundary impedance, defined as the spectral ratio between traction and velocity at the boundary node,

$$Z(\omega) = -\tau/i\omega d_0, \tag{62}$$

is obtained by combining the dispersion relation Equation 61 with Equation 59:

$$Z(\omega) = \sqrt{G_1^2 - G_2^2}/\omega \tag{63}$$

Inhomogeneous wave modes (exponential spatial decay) appear when the wavenumber k takes a non-vanishing imaginary part. This is possible only within specific stopping bands. The edges of the stopping bands are defined by the condition $\cos(kh) = \pm 1$, which occurs either at the crossing points $G_2 = \pm G_1$ or at the common poles of G_1 and G_2. These two possibilities correspond respectively to zeros and poles of $Z^2(\omega)$.