Dynamic wedge failure reveals anomalous energy radiation of shallow subduction earthquakes

Shuo Ma\textsuperscript{a,}\textsuperscript{*}, Evan T. Hirakawa\textsuperscript{a,}\textsuperscript{b}

\textsuperscript{a} Department of Geological Sciences, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-1020, USA
\textsuperscript{b} Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0225, USA

\textbf{ARTICLE INFO}

Article history:
Received 29 December 2012
Received in revised form 8 May 2013
Accepted 10 May 2013
Editor: P. Shearer
Available online 5 June 2013

\textbf{Keywords:}
Coulomb wedge theory
dynamic pore pressure change
rupture dynamics
 tsunami generation
anomalous seismic radiation
shallow subduction zone

\textbf{ABSTRACT}

Dynamically induced Coulomb failure in the overriding wedge significantly affects energy radiation of shallow subduction earthquakes. For a wedge on the verge of failure, extensive fluid-assisted coseismic failure due to updip rupture causes significant seafloor uplift above a shallow dipping basal fault. The large inelastic uplift, greatly enhanced by the presence of free surface, significantly dilates the fault behind the rupture front during the rupture propagation, which reduces the effective normal stress and sliding friction on the fault, and increases the dynamic stress drop and slip velocity. As a result, slip-velocity time histories in the shallow section of the fault tend to have a 'snail-like' shape, leading to depletion of high frequencies in the slip velocity field and the resultant source time function. We also show that the failure in the wedge acts as a large energy sink (while contributing to seismic moment), giving rise to distributed heat generation (i.e., small heat flow anomaly across the fault), moment-scaled radiated energy, slow rupture velocity, and small directivity, which provides a unifying interpretation for nearly all anomalous observations documented for shallow subduction earthquakes.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The shallow reaches of subduction interfaces (upper 10–15 km) host earthquakes of anomalous energy release attributes. Earthquake ruptures there can be exceptionally tsunamigenic, but produce weak high-frequency seismic radiation (low body- and surface-wave magnitudes), as evidenced in tsunami earthquakes (e.g., Kanamori, 1972; Pelayo and Wiens, 1992; Polet and Kanamori, 2000; Ammon et al., 2006; Kanamori et al., 2010; Newman et al., 2011; Bilek et al., 2011). Recent large tsunamiogenic events, such as the 2004 $M_W$ 9.2 Sumatra earthquake and 2011 $M_W$ 9.0 Tohoku-Oki earthquake, are not tsunami earthquakes because these earthquakes also radiated significant amount of high-frequency energy from the deeper section of the plate interface. However, their rupture characteristics in the upper 10–15 km are very similar to those of tsunami earthquakes (Lay et al., 2012; Yao et al., 2013).

Numerous observations (e.g., Lay and Bilek, 2007) indicate that these shallow ruptures are also associated with unusually long rupture duration, slow rupture velocity, small stress drop, as well as low moment-scaled radiated energy (Newman and Okal, 1998; Lay et al., 2012). What gives rise to these anomalous characteristics in the shallow subduction zone and how they relate to large tsunamigenesis are, however, still not well understood.

The prevailing hypothesis for the long rupture duration and slow rupture velocity is the presence of a thin layer of subducted sediments in the fault zone (e.g., Polet and Kanamori, 2000). Lower rigidity of sediments also leads to larger slip if seismic moment is kept constant, which causes larger seafloor displacement and accounts for tsunami generation. However, dynamic rupture models of earthquakes incorporating a low-velocity fault zone with elastic material response (e.g., Harris and Day, 1997; Huang and Ampuero, 2011) disfavor this hypothesis. Dynamic models show that the low-velocity fault zone does not necessarily lead to slow rupture velocity; the rupture velocity is mostly controlled by the host country rock. More importantly, trapped energy in the fault zone makes the rupture more ‘pulse-like’ and can instead lead to supershear rupture propagation, which inevitably enhances high-frequency radiation and violates the observations.

The horst-and-graben structures on the subducting plate have also been suggested to generate a strong contact on the fault surface to allow the rupture to reach the trench (Polet and Kanamori, 2000). The rough fault surface also leads to large fracture energy (Venkataraman and Kanamori, 2004), contributing to slow rupture propagation. This hypothesis, however, violates again the physics that fault roughness tends to increase, rather than decrease the high-frequency radiation (e.g., Dunham et al., 2011; Shi and Day, 2013).

Another puzzle in the shallow subduction zone relates to the mechanism of tsunami generation. Most current work attributes
tsunami generation to large slip near the trench (e.g., Satake and Tanioka, 1999; Fujii and Satake, 2007; Fujii et al., 2011; Lay et al., 2012). However, large slip on a shallow-dipping plate interface (dip less than 15°, as is the case in most subduction zones), causes predominantly horizontal displacement, which is inefficient to generate tsunamis. In an elastic model the seafloor uplift near the trench \( U_T \) can be approximated by \( U_T \approx \Delta \sin \theta \), where \( \Delta \) is the slip and \( \theta \) is the fault dip. For a very shallow fault dip, finite seafloor uplift would inevitably require large slip. This relationship also predicts that a shallower fault dip leads to larger fault slip, which is why shallow fault dips are favored in reconciling the discrepancies between the moment magnitude and the magnitude estimated from tsunami data for tsunami earthquakes (see the review by Lay and Bilek, 2007). The physics that a shallower fault dip leads to larger slip is unclear. In fact, for a very shallow fault dip, this relationship is expected to break down, because the shallow-tapered geometry of wedge can lead to large stress concentrations, which causes the wedge to fail and makes this mechanism likely implausible.

Motivated by the critical taper theory for accretionary wedges (e.g., Davis et al., 1983; Dahlen, 1990), Ma (2012) proposed that for a wedge on the verge of failure, the pore pressure increase due to updip rupture propagation significantly weakens the wedge, leading to extensive Coulomb failure (inelastic deformation) in the wedge. Extensive failure gives rise to slow rupture velocity and large seafloor uplift updward from the trench, which is greatly enhanced by the shallow fault dip. The peak seafloor uplift occurs above the fault section where slip is tapered to nearly zero, i.e., large near-trench slip is not required to account for large tsunami generation. This self-consistent mechanism solves many of the aforementioned paradoxes.

In retrospect, we found that Seno (2000) and Tanioka and Seno (2001) proposed a similar mechanism for tsunami generation. Their work, however, did not include full dynamic analyses of off-fault failure on rupture propagation and tsunami generation. Also the failure is localized near the toe at the trench, compared to the extensive failure in Ma (2012) and this study.

Here we extend this wedge-failure mechanism to explain the anomalous energy radiation attributes. We show that large seafloor uplift due to extensive wedge failure tends to produce ‘snail-like’ slip-velocity time histories on a shallow-dipping plate interface, which greatly increases the low-frequency content in the slip-velocity field and leads to deficiencies in the high-frequency radiation. Large energy dissipation due to the inelastic deformation also explains slow rupture velocity and low moment-scaled radiated energy. Efficient tsunami generation, deficiency in the high-frequency radiation, slow rupture velocity, and low moment-scaled radiated energy are thus all consistent manifestations of extensive coseismic failure in the wedge, which is a unifying interpretation.

Slip near the trench is likely small due to extensive wedge failure, producing small horizontal seafloor displacement near the trench. The large horizontal seafloor displacement observed/inferred for the 2011 Tohoku-Oki earthquake (e.g., Fujiiwara et al., 2011) was attributed to large slip at the trench (e.g., Lay et al., 2012). We suggest that it could also be due to slope instability (landslides) induced by seafloor uplift in this mechanism. Because the wedge is on the verge of failure initially coseismic wedge deformation can easily trigger the landslides. The coupled effect of dynamic rupture propagation, wedge failure, seafloor uplift, landslides, and tsunami propagation will be addressed in a forthcoming paper.

2. Emergence of ‘snail-like’ rupture and radiation characteristics

We use a 2D plane strain model (Fig. 1), similar to Ma (2012). The stresses in the upper 15.53 km of the overriding wedge are subject to the Mohr–Coulomb failure criterion. Dynamic pore pressure change in an undrained condition is incorporated. The model is based on the poroelastoplastic model of Viesca et al. (2008) without dilatancy and hardening, and uses a slip-weakening friction criterion on the fault. The initial stresses in the medium linearly increase with depth, so does the cohesion (e.g., Wang and Hu, 2006). The model parameters are listed in Table 1. The main difference between these parameters and those of Ma (2012) is the reduction of the static frictional coefficient (\( \mu_s \)) from 0.85 to 0.6 and the Skempton's coefficient (\( B \)) from 0.7 to 0.3. The dynamic stress drop is also reduced to 1% of the normal stress on the fault if normal stress variations are ignored (\( \mu_0-\mu_d = 0.01 \), where \( \mu_0 \) is the ratio of the initial shear stress to the normal stress on the fault and \( \mu_d \) is the dynamic frictional coefficient). Small Skempton's coefficients (0.1–0.6) were suggested for subduction zones (e.g., Wang and Davis, 1996). Other parameters in the model are accordingly adjusted to give rise to a wedge close to failure. In this work we use a larger closeness-to-failure (CF) value of 0.99. The CF is defined as \( CF = \frac{\rho_0}{\rho_{max}} \), where \( \rho_{max} \) is the maximum initial shear stress at a point, strength parameters \( \beta \) and \( \mu \) are functions of cohesion \( c \) and internal frictional angle \( \phi \) as \( \beta = c \cos \phi \) and \( \mu = \sin \phi \), and the initial mean effective normal stress is \( \sigma_{0m}^n = \frac{1}{2}(\sigma_{01}^n + \sigma_{02}^n) \). In the undrained condition, undrained strength parameters should be used to define \( CF: \beta_0 = \beta + \beta_0 \sigma_{01}^n \) and \( \mu_0 = (1 - B) \mu \), which essentially gives the same CF value as in the drained case. Note that here we corrected a minor error in the equation for \( \beta_0 \) in Ma (2012), which made the wedge in the undrained cases of that work closer to failure than the value specified (\( CF=0.9 \)).

As we will see, there is a tradeoff among Skempton's coefficient \( B \), stress drop, and CF. The combination of small \( B \) and stress drop, and large CF values can also give rise to extensive failure in the wedge. Results will not be sensitive to the Skempton's coefficient if the wedge is on the verge of failure.

We carry out elastic and inelastic simulations using the same parameters listed in Table 1, except that in the elastic simulation an infinite cohesion is used such that no failure occurs in the wedge. We show results below for a hydrostatic case (\( \lambda_0 = 0.3745 \), where \( \lambda_0 \) is the ratio of initial pore pressure to the lithostatic stress and is a constant in this work) unless otherwise noted. The same pore pressure ratio is used on the fault as in the wedge.

As was shown in Ma (2012), the pore pressure increases due to updip rupture and the nearly critical state of the wedge cause extensive failure in the wedge (Figs. 2, A1 and Movie in the...
The failure gives rise to slow rupture velocity at the shallow section of the fault and large sea floor uplift landward from trench. The total seismic moment in the inelastic simulation (16% contributed by off-fault failure) is approximately 21% less than that in the elastic simulation (1.99 × 10^16 N m vs. 2.52 × 10^16 N m); however, it produces larger static peak sea floor uplift (3.17 m vs. 2.74 m). The dynamic sea floor uplift during the rupture propagation is in fact larger than the static uplift in both cases (Fig. A1). The peak uplift reaches 4.12 m at 27.46 s (after the rupture breaks the trench) in the elastic case, compared to 3.95 m at 25.86 s (before the rupture breaks the trench) in the inelastic case. The timing and location differences of dynamic peak uplift between these cases should provide detectable differences in the tsunami data. The larger dynamic peak uplift in the elastic case is partly due to the slightly larger fault dip (15°) we chose. For a shallower fault dip, the dynamic peak uplift in the elastic case is expected to be smaller because the shallower fault dip decreases the elastic uplift, but promotes more inelastic uplift. When we compare the efficiency of generating dynamic sea floor uplift (peak dynamic uplift normalized by seismic moment) the inelastic case is still 21.3% more efficient to generate the uplift.

Supplementary material related to this article can be found online at http://doi:10.1016/j.ultramic.2010.09.003.

In the elastic case, we see supershear rupture velocity near the trench, strong rupture breakout of the sea floor (generating large reflections), and large slip at the trench. In contrast, in the inelastic case, the rupture velocity is slow (approximately 2 km/s) near the trench, rupture breaks the trench gently, and the slip is tapered to nearly zero toward the trench. As a result, the sea floor displacement is nearly continuous in the inelastic case, compared to the large discontinuity at the trench in the elastic case. Supershear rupture velocity, strong rupture breakout, and large slip velocity near the trench in the elastic case would produce large high-frequency radiation and strong rupture directivity, which is inconsistent with the observations we summarized earlier for shallow subduction earthquakes.

Table 1
Calculation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial effective stresses in medium (Positive in tension)</td>
<td>$\bar{s}_3 = \bar{s}_z = -(1-\lambda_w)\rho g z$, $\bar{s}_1 = \bar{s}_x = 3.7469\rho g z$</td>
</tr>
<tr>
<td>Initial friction ($\mu_0$)</td>
<td>0.58</td>
</tr>
<tr>
<td>Pore pressure to lithostatic stress ratio ($\lambda_w$)</td>
<td>0.3745, 0.8</td>
</tr>
<tr>
<td>Static friction ($\mu_s$)</td>
<td>0.6</td>
</tr>
<tr>
<td>Dynamic friction ($\mu_d$)</td>
<td>0.57</td>
</tr>
<tr>
<td>Slip-weakening distance ($D_c$)</td>
<td>0.2 m</td>
</tr>
<tr>
<td>Cohesion ($c$)</td>
<td>0.017(\rho g z)</td>
</tr>
<tr>
<td>Internal friction (\tan $\phi$)</td>
<td>0.7095</td>
</tr>
<tr>
<td>Closeness-to-Failure (CF)</td>
<td>0.99</td>
</tr>
<tr>
<td>Skempton’s coefficient ($B$)</td>
<td>0.3</td>
</tr>
<tr>
<td>Undrained Poisson's ratio ($\nu_u$)</td>
<td>0.2788</td>
</tr>
<tr>
<td>Element size on fault</td>
<td>100 m</td>
</tr>
<tr>
<td>Time step</td>
<td>0.0025 s</td>
</tr>
<tr>
<td>Calculation time</td>
<td>100 s</td>
</tr>
</tbody>
</table>
A pronounced feature behind the rupture front in the slip-velocity field in the inelastic case is particularly noteworthy, which is only present at shallow depths (Fig. 2d). A similar feature was also shown in Ma (2012, Fig. 3). In the elastic case (Fig. 2c), this feature is present, but much less pronounced. To find out what this feature is, we plot the slip velocity time histories at 25 km and 65 km down dip on the fault (Fig. 3a and b). At 65 km down dip, the slip velocity functions are ‘crack-like’—slip velocity decays sharply after initiation and takes a long time to heal (slip duration is comparable to rupture duration). No inelastic deformation occurs around this point. The two slip velocity functions and their frequency content are nearly identical, except for the reflections seen in the elastic case (Fig. 3a). At 25 km down dip, we see distinctly different characteristics (Fig. 3b). The slip velocity decays much more gradually than at depth although durations are shorter. Due to the gradual decay the frequency content is more depleted in the high frequency range above ~0.1 Hz at this shallow point (Fig. 3c). Furthermore, when we look at the shape of the slip velocity function in the inelastic case, the slip velocity decays first, then rises again over a long duration forming a large ‘bump’. The peak slip velocity is not at the slip initiation, but on the ‘bump’. The overall shape of the slip velocity function resembles a ‘snail’. The initial slip velocity increase and the ‘bump’ correspond to the ‘snail’s antenna and shell’, respectively. The pronounced feature seen in Fig. 2d thus represents ‘snails’ shells’ of slip velocity histories at points on the shallow section of the fault. The elastic and inelastic slip durations (rise times) at this point (14.99 s and 17.11 s) are comparable. However, the overall smoothness and the ‘shell’ of the ‘snail-like’ slip velocity function lead to further depletion in frequencies above ~0.1 Hz (Fig. 3c).

The spatial distribution of slip velocity on the fault is also ‘snail-like’ (Fig. 3e), where the peak slip velocity is not at the rupture front, but rather behind it. We also note a ‘snail-like’ shape of the deforming seafloor in the inelastic case, propagating horizontally toward the trench as the rupture propagates updip (Fig. 3e).

What gives rise to ‘snail-like’ rupture characteristics at shallow depths in the inelastic case, but not in the elastic case? A strong correlation is clearly seen between the increase in slip velocity behind the rupture front and large induced seafloor uplift (Fig. 3e). To confirm this correlation, we plot the normal stress change at 25 km down dip (Fig. 3f). A more gradual and longer-lasting effective normal stress reduction is seen after slip initiates in the inelastic case than the elastic case, which is consistent with the dilation induced by the seafloor uplift. The seafloor dilation causes a reduction in the total stress, which counteracts the effect of reducing pore pressure leading to the effective normal stress reduction on the fault. The nearly vertical seafloor displacement in the inelastic case exerts a more direct and prolonged dilation to the shallow-dipping fault than the nearly horizontal displacement in the elastic case (Fig. 3d and e). In both cases, the effective normal stress reduction increases the shear stress drop (Fig. 3g).

Fig. 3. (a and b) Time histories of slip velocity at 2 points on the fault (25 km and 65 down dip). Blue curves are the elastic case and red curves the inelastic case. The spectra are shown in (c). The slip velocity functions at shallower depth are more depleted in high frequencies. The ‘snail-like’ slip velocity in the inelastic case leads to further reduction in the high-frequency content. (d and e) Snapshots of seafloor displacement (arrows), and slip rate on fault (magenta curves), and seismic potency density (color scale) at a section of the fault. The elastic and inelastic slip durations (rise times) at this point (14.99 s and 17.11 s) are comparable. However, the overall smoothness and the ‘shell’ of the ‘snail-like’ slip velocity function lead to further depletion in frequencies above ~0.1 Hz (Fig. 3c).
which further increases the slip velocity, such that the slip-velocity functions have the more gradual decrease (Fig. 3b). However, only in the inelastic case do we see the more significant seafloor dilation behind the rupture front, which increases the slip velocity to larger amplitude than at the slip initiation, forming the ‘snail-like’ slip velocity function. We point out that even if the slip-velocity function is not ‘snail-like’ (i.e., the peak is still around the initiation time), the tendency of the inelastic failure and free surface is to induce a ‘bump’ to enhance the low-frequency radiation.

The ‘snail-like’ slip velocity time histories directly lead to a much smoother moment-rate time function (including both on-and off-fault moments) in the inelastic case than the elastic case, which gives rise to depletion of energy in nearly the entire frequency band above 0.01 Hz compared to the elastic case (Fig. 4). Extensive coseismic wedge failure and ‘snail-like’ rupture characteristics thus provide a rigorous physical explanation for the well-documented deficiency in high-frequency radiation of shallow subduction earthquakes.

3. Radiated energy

We also calculate the total radiated energy using the stress and strain fields in the simulations. In the elastic case, the calculation is straightforward (e.g., Kostrov, 1974; Ma and Archuleta, 2006). However, no equations are readily available in the inelastic case because the work done by off-fault failure needs to be considered in the total energy budget.

Here we derive the energy equation for off-fault failure based on physical principles of mechanical work. In the simulations, when the stress state at a point violates the Mohr–Coulomb failure criterion in the inelastic simulation we reduce all the deviatoric stresses ($\sigma_{ij}$) by a common factor to meet the failure criterion (e.g., Andrews, 2005a). No adjustment is applied to the mean stress. The stress adjustments, called stress glut (Backus and Mulcahy, 1976), are sources of seismic radiation and they represent increments to the volume density of seismic moment tensor $dm_{ij}/dm_{0} = c_{ijkl} \Delta \varepsilon_{kl}^{ij}$, where $c_{ijkl}$ is the linear elastic tensor and $\Delta \varepsilon_{kl}^{ij}$ is the inelastic strain increment. In the simulations, we use $dm_{ij} = 2Gd\varepsilon_{ij}^{0}$ ($G$ is the shear modulus) and $dm_{0} = 0$, which is a special case. The moment sources, representing the failure off the fault, are applied to an otherwise elastic medium. The net work done by off-fault moment sources during the earthquake can thus be written as

$$W_{\text{off}} = \int_{V} w_{\text{off}} \, dV = \int_{V} \int_{0}^{t_{f}} m_{ij}(t) \varepsilon_{ij}^{0}(t) \, dt \, dV - \frac{1}{2} \int_{V} m_{ij} \varepsilon_{ij}^{0} \, dV,$$  \hspace{1cm} (1)

where $w_{\text{off}}$ is the volume density of the net work done off the fault, $\varepsilon_{ij}$ is the strain tensor, superscripts 0 and 1 represent the initial and final states, respectively, and overdot denotes the time derivative. We take the initial strain field $\varepsilon_{ij}^{0}$ to be zero and the initial stress field $\sigma_{ij}^{0}$ nonzero. Integrating Eq. (1) by parts, we obtain

$$W_{\text{off}} = \int_{V} \{m_{ij}(t)\varepsilon_{ij}^{0}(t)\}_{0}^{t_{f}} \, dV - \int_{V} \int_{0}^{t_{f}} m_{ij}(t)\varepsilon_{ij}^{0}(t) \, dt \, dV - \frac{1}{2} \int_{V} m_{ij} \varepsilon_{ij}^{0} \, dV$$

$$= \frac{1}{2} \int_{V} m_{ij} \varepsilon_{ij}^{0} \, dV - \int_{V} \int_{0}^{t_{f}} m_{ij}(t)\varepsilon_{ij}^{0}(t) \, dt \, dV$$

$$= \int_{V} \frac{1}{2} c_{ijkl} \Delta \varepsilon_{kl}^{ij} \varepsilon_{ij}^{0} \, dV - \int_{V} \int_{0}^{t_{f}} c_{ijkl} \Delta \varepsilon_{kl}^{ij}(t)\varepsilon_{ij}^{0}(t) \, dt \, dV.$$  

Introducing Hooke’s law ($\sigma_{ij} - \sigma_{ij}^{0} = c_{ijkl} \varepsilon_{kl}^{ij}$) and using the symmetry of $c_{ijkl} = c_{klij}$, we have

$$W_{\text{off}} = \int_{V} \frac{1}{2} (\sigma_{ij}^{0} + \sigma_{ij}^{0}) \varepsilon_{ij}^{0} \, dV - \int_{V} \int_{0}^{t_{f}} \sigma_{ij}(t)\varepsilon_{ij}^{0}(t) \, dt \, dV,$$

which can be finally written as

$$W_{\text{off}} = \int_{V} \frac{1}{2} (\sigma_{ij}^{0} + \sigma_{ij}^{0}) \varepsilon_{ij}^{0} \, dV - \int_{V} \int_{0}^{t_{f}} \sigma_{ij}(t)\varepsilon_{ij}^{0}(t) \, dt \, dV.$$  \hspace{1cm} (2)

The first term in Eq. (2) can be shown to be part of the elastic strain energy and gravitational potential energy changes in the medium, and the second term is the energy dissipation by the inelastic deformation. In the results below, we calculate the net work done off the fault using Eq. (1) for its easy implementation.

The moment due to off-fault failure can be calculated as follows. At a point, the seismic potency density as a function of time is

$$\eta(t) = \int_{V} d\eta,$$  \hspace{1cm} (3)

where the increment due to stress glut is $d\eta = \sqrt{1/2} \, dm_{ij} / \sqrt{G}$. In 2D, we can also write this term explicitly as $d\eta = \sqrt{(dm_{xx} - dm_{yy})^2 + (dm_{xy})^2 / G}$. In previous works (Ma, 2008, 2009, 2012; Ma and Andrews, 2010) the quantity $\eta$ was termed the accumulated inelastic strain. The total off-fault moment can thus be obtained by $M_{\text{off}} = \int_{V} \sigma_{ij}(t)\varepsilon_{ij}^{0}(t) \, dt \, dV$.

The net work done on the fault $W_{\text{on}}$ in both elastic and inelastic cases can be calculated by

$$W_{\text{on}} = \frac{1}{2} (\sigma_{ij}^{0} + \sigma_{ij}^{0}) \Delta \varepsilon_{ij}^{0} \eta_{ij} \Delta \varepsilon \Delta \sigma - \frac{1}{2} \int_{V} \int_{0}^{t_{f}} \sigma_{ij}(t)\Delta \varepsilon_{ij}^{0}(t) \eta_{ij} \, dt \, d\Sigma,$$  \hspace{1cm} (4)

where $\Delta \varepsilon_{ij}$ is the fault slip and $\eta_{ij}$ is the unit outer normal of the fault plane $\Sigma$. This equation is same as Eq. (5) of Ma and Archuleta
The total radiated energy is thus

\[ P_{\text{source}} = W_{\text{on}} + W_{\text{off}}. \]

We also calculate the work done by stress changes due to seismic waves going through a 150 km radius semicircle (surface \( \omega \)) enclosing the fault and the off-fault failure zone, \( P_{\text{circle}} \), using

\[ P_{\text{circle}} = \int_{\omega} w_{\text{circle}} \, dS = \int_{\omega} \int_0^{t_{\text{ij}}} -\sigma_{ij}(t) - \sigma_{ij}^0 \, u_{ij} \, dt \, dS \\
+ \int_{\omega} \frac{1}{2} \left( \varepsilon_{xy}(t) - \varepsilon_{xy}^0 \right) u_{xy} \, dt \, dS,
\]

where \( w_{\text{circle}} \) is the density of the net work done on the surface \( S \), \( u_{ij} \) is the displacement field, \( \varepsilon_{xy} \) is the strain tensor of the inelastic case (1.4610^-10), and the two terms on the right side are, respectively, the total work (\( W_{\text{total}} \)) and static work (\( W_{\text{static}} \)) done by stress changes on the surface \( S \). This equation is equivalent to Eqs. (7) and (8) of Ma and Archuleta (2006). The conservation of energy requires

\[ P_{\text{source}} = P_{\text{circle}}. \]

The failure in the wedge is seen to act as a large energy sink, while contributing to the total moment, which gives rise to a lower moment-scaled radiated energy (Fig. 5). The net energy loss in the wedge (\( W_{\text{off}} \)) is approximately 50% of the net work done on the fault (\( W_{\text{on}} \)) in both the hydrostatic case and an overpressure case (\( \omega = 0.8 \)), which is thus approximately equal to the total radiated energy of the earthquake. Such large net energy loss in the wedge would be manifested by distributed heat generation, i.e., small heat flow anomaly across the basal fault. The off-fault failure gives rise to a 40% reduction of moment-scaled radiated energy in the hydrostatic case (1.46 × 10^-5 vs. 2.43 × 10^-7) and a 45% reduction in the overpressure case (4.22 × 10^-5 vs. 7.70 × 10^-8) when compared to the elastic cases. Shallow subduction events can have moment-scaled radiated energy an order or more lower than regular earthquakes (Newman and Okal, 1998). Considering different environments in which these earthquakes occur, we compare the moment-scaled radiated energy in the elastic hydrostatic case with the inelastic pressure case, obtaining a factor of 5.8 differences. A shallow fault dip, as will be discussed below, would further lower the moment-scaled radiated energy. Possible heterogeneities in the shallow subduction zone (e.g., material properties, stress drop, and pore pressure) and inelastic volumetric deformation (e.g., dilatancy), not included in our model, may also help explain the discrepancies. In Fig. 5, we also see smaller rupture directivity due to large energy dissipation in the wedge, as illustrated by a more uniform distribution of work density on the semicircle in the inelastic cases.

Here we give an interpretation to slow rupture velocity in the inelastic case from an energy perspective. The slow rupture velocity is caused by the large energy dissipation in the wedge, similar to the slowness of a snail due to large energy expended in carrying its shell. The rupture propagation with extensive failure is thus ‘snail-like’ also in the energy sense.

4. Discussion

4.1. The Importance of free surface

The presence of free surface significantly affects the extent and magnitude of Coulomb failure when the failure criterion is pressure-dependent, giving rise to ‘flower-like’ fault zone structures (Ma, 2008, 2009; Ma and Andrews, 2010). In the shallow subduction zone, for a wedge on the verge of failure above a shallow-dipping fault (e.g., Dahlen, 1990), the free surface greatly promotes the failure in the wedge, especially at shallow depths. The extensive failure leads to anomalous observations in the shallow subduction zone, as demonstrated here. When the fault dip is shallower, the free surface-enhanced inelastic effects will be more significant because the lower confining pressure in the wedge makes the material weaker. The dilatational effect of the seafloor uplift (Fig. 3e) will also be more pronounced. Therefore we expect to see a slower rupture velocity, more depletion in the high-frequency radiation, more efficient tsunami generation, and a lower moment-scaled radiated energy for earthquakes on a shallower dipping fault. Due to significant effects of the free surface, whole-space models without the free surface (e.g., Noda and Lapusta, 2013) probably do not capture the essential physics of shallow subduction earthquakes.

Lay et al. (2012) and Yao et al. (2013) reported that rupture characteristics in subduction zones are depth-dependent; deep ruptures radiate more high-frequency energy while shallow ruptures are more depleted in high frequencies. These observations could also be explained by the free surface effect. At shallow depths the free surface causes much larger motion on the hanging wall than the footwall (e.g., Oglesby et al., 1998; Shi et al., 1998; Ma and Beroza, 2008). The larger hanging wall motion greatly dilates the fault and reduces the effective normal stress on the fault. The normal stress reduction causes the slip velocity functions at shallow depths to decay more gradually than at deeper depths. So even in the elastic case we expect shallow ruptures to be more deficient in high frequencies (Fig. 3a–c). When we consider the free surface-enhanced wedge failure, shallow subduction earthquake ruptures would be even more depleted in high-frequency radiation than deep ruptures (Fig. 3c), which well explains the unusual radiation characteristics of Domain A earthquakes documented in Lay et al. (2012).

Our mechanism also provides an alternative interpretation to the structural complexity differences between the inner and outer wedges reported in Wang and Hu (2006). Wang and Hu (2006) proposed that coseismic strengthening of friction on the shallow section of the plate interface could explain more intense deformation of the outer wedge. In our model, the velocity-strengthening friction is not used. Although the wedge is everywhere close to failure initially (\( CF = 0.99 \)), we see more intense failure at shallow depths (outer wedge), which is essentially due to a closer proximity to the seafloor. Despite possibly an end-member case, our models show that coseismic wedge failure alone could explain the structural complexity differences.

4.2. Shallow fault dip, long rupture duration, and location of peak seafloor uplift

Actual subduction zones have fault dips shallower than the 15° modeled here (e.g., Davis et al., 1983). A shallower fault dip not only significantly increases the fault width, but also lowers the confining pressure in the wedge, as mentioned earlier. The larger fault area being closer to the free surface would greatly promote coseismic failure in the wedge, leading to a slower rupture velocity. The larger fault width and a slower rupture velocity together would give rise to longer rupture duration than the one shown here (~40 s). Rupture propagation along strike, not included in our model, is also expected to increase the rupture duration. Both may contribute to the ~100 s long rupture duration of tsunami earthquakes (e.g., Pote and Kanamori, 2000).

Due to the more significant wedge failure for a shallower fault dip, the peak seafloor uplift is likely more landward from the trench. A broader area of the seafloor will be predominantly uplifted if the rupture propagates close to the trench. Observations in the 1964 Alaska earthquake (Palmer, 1969) seem to support this scenario. The fault dip of the Alaska earthquake was as shallow as 3°–4° (Brocher et al., 1994). The peak uplift of 11.3 m occurred close to the shore on Tungul Island; the seaward zone on the seafloor could be uplifted even higher. Middle Island, ~90 km further seaward to
the trench, was also mainly uplifted by 3.5 m. Both islands did not experience much horizontal displacement (Plafker, 1969), unlike the observations in the 2011 Tohoku–Oki earthquake (e.g., Fujiwara et al., 2011). The uplifts on these islands were associated with secondary splay faulting; however, both the hanging wall and footwall of the splay fault (Patton Bay fault) on Montague Island were uplifted by more than 5 m. This uplift along with the broad regional uplift during this earthquake can be well explained by our mechanism. More supporting observations for this mechanism can be found in the 1896 Sanriku earthquake (Tanioka and Seno, 2001) and the 2010 Mentawai earthquake (Hill et al., 2012).

4.3. Sediments in the fault zone and fault roughness

Although subducted sediments and horst and graben structures on the plate interface (e.g., Polet and Kanamori, 2000) were proposed to explain the long rupture duration and slow rupture velocity, they tend to generate more high-frequency radiation if the material response is elastic, as was pointed out earlier. Our model assumes a homogeneous half space and a planar fault. Slow rupture velocity and long rupture duration are results of extensive inelastic wedge failure. It would be incorrect to attribute the slow rupture velocity and long rupture duration only to the presence of sediments. Effects of sediments and fault roughness can potentially be important, but inelastic mechanisms like the one presented here are needed to suppress the high-frequency generation. These mechanisms can interact with each other and all may play important roles in the dynamics of shallow subduction earthquakes.

4.4. Critical taper wedge and frictional stability on the basal fault

No velocity-strengthening friction is used in our model. However, our inelastic models see velocity-strengthening friction behaviors, such as small slip and slow rupture velocity near the trench. In the extreme overpressure case, off-fault failure can stall the rupture and even prevent the rupture from reaching the trench (see Fig. 3f in Ma, 2012). Coseismic wedge failure is clearly not a frictional property on the fault.

In most current work, fault zone complexities are incorporated in a rate- and state-variable friction framework to explain the stability of subduction megathrusts (e.g., Scholz, 1998). Off-fault material is often assumed elastic although coseismic wedge weakening was clearly recognized (e.g., Wang and Hu, 2006).
Fig. A1. Snapshots of seafloor displacement (arrows), slip velocity on fault (magenta curves), and seismic potency density (color scale) in the wedge are shown for the elastic and inelastic simulations. Pore pressure is initially hydrostatic. Red and green arrows show displacement on the hanging wall and footwall, respectively. The number above each panel indicates the maximum seafloor uplift at that time. Note that the seafloor uplift during the rupture propagation is larger than the final static uplift, which may significantly affect tsunami generation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
In doing so, all possible complexities off the fault are mapped onto a single fault surface, which could be a serious oversimplification and can lead to extreme frictional conditions on the fault (e.g., Noda and Lapusta, 2013). We suggest that the velocity-strengthening friction on the fault could be a proxy to extensive failure off the fault, but probably does not fully represent the off-fault deformation mechanism.

We quote an important result of the critical taper theory from Davis et al. (1983): “A wedge of less than critical taper will not slide when pushed but will deform internally, steepening its surface slope until the critical taper is attained…. rock deforms until the wedge attains a steady state or critical taper and then slides stably, continuing to grow self-similarly as additional material is accreted at the toe.” Clearly the stress state in the wedge significantly controls the stability of basal fault. The general consensus about subduction megathrusts in the upper 10–15 km being stable (a velocity-strengthening region) could in fact be indicative of a critically or supercritically stressed overriding wedge.

Much attention is focused on the fault zone properties in controlling earthquake characteristics in subduction zones (e.g., Byrne et al., 1988; Scholz, 1998; Lay et al., 2012). We suggest that the critical or nearly critical state of overriding wedges (e.g., Davis et al., 1983; Dahlen, 1990; Wang and Hu, 2006) should be taken into account and potential role of wedge failure could change our current understanding of deformation processes in shallow subduction zones.

4.5. Possible inelastic effects in the 1999 Chi-Chi earthquake

Seno (2000) noted abnormal uplift near a riverbed in the northern rupture zone of the 1999 Chi-Chi earthquake. There were also striking differences in ground motion recorded in northern and southern sections and ground motions in the north are much more depleted in high-frequencies than in the south (e.g., Ma et al., 2003). These characteristics are very similar to those of a shallow subduction earthquake discussed in this paper. We propose that extensive off-fault failure is responsible for the enhanced long-period radiation in the northern section, compared to other proposed mechanisms, such as fault lubrication (Ma et al., 2003) and thermal pressurization (Andrews, 2005b; Noda and Lapusta, 2013). Well-recorded seismic and geodetic data during the Chi-Chi earthquake can be used to test this hypothesis.

5. Conclusions

We show that extensive coseismic failure in the wedge causes significant vertical seafloor displacement above a shallow dipping fault that strongly dilates the fault behind the rupture front, promoting ‘snail-like’ rupture characteristics. The greatly enhanced low-frequency content of the slip velocity field leads to depletion in the high-frequency radiation. We also demonstrate that extensive failure in the wedge is a large energy sink, generating distributed heat and small heat flow anomaly, which gives rise to low moment-scaled radiated energy, slow rupture velocity, and small directivity. These effects will be more significant as the fault dip shallows. This mechanism thus provides a unifying interpretation for nearly all anomalous observations documented for shallow subduction earthquakes and may solve the 40-yr old puzzle about tsunami earthquakes (Kanamori, 1972).

The failure in the wedge likely plays an important role in the stress accumulation and relaxation during earthquake cycles in the shallow subduction zone, which needs to be incorporated in current coseismic, postseismic and interseismic deformation models. Elastic Green’s functions can lead to large biases in slip models (Yao and Ma, 2012), as can be expected from the distinctly different displacement fields in the elastic and inelastic simulations (Figs. 2 and A1).

Acknowledgments

The first author wishes to thank Joe Andrews for his encouragement and support of this work; the derivation of the inelastic energy equations was motivated by a discussion with him at the 2012 SSA meeting. We also thank Eric Dunham, an anonymous reviewer, and Peter Shearer (Editor), for their insightful reviews and suggestions, which dramatically improved the clarity of the paper. This work was supported by NSF Grant EAR-1045369. The simulations used Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by NSF Grant OCI-105375 using allocations of the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0529922 and USGS Cooperative Agreement 07HQAG0008.

Appendix A. Rupture snapshots of the elastic and inelastic simulations for the hydrostatic case

See Fig. A1.

References


Hill, E.M., et al., 2012. The 2010 Mw 7.8 Mentawai earthquake: very shallow source of a rare tsunami earthquake determined from tsunami field survey and near-
Lay, T.H., Kanamori, C.A., Ammon, K.D., Koper, A.R., Hutko, L., Ye, H., Yue, Rushing, T.M.,
Ma, S., 2012. A self-consistent mechanism for slow dynamic deformation and
tsunami generation for earthquakes in the shallow subduction zone. Geophys.
Kostrov, B.V., 1974. Seismic moment and energy of earthquakes, and seismic
122
Ma, S., Andrews, D.J., 2010. Inelastic off-fault response and three-dimensional
Ma, S., Archuleta, R.J., 2006. Radiated seismic energy based on dynamic rupture
Ma, S., Beroza, G.C., 2008. Rupture dynamics on a bi-material interface for dipping
Shi, Z., Day, S.M., 2013. Rupture dynamics and ground motion from 3-D rough-fault
Seno, T., 2000. The 21 September, 1999 Chi-Chi earthquake in Taiwan: implications
Shi, Z., Day, S.M., 2013. Rupture dynamics and ground motion from 3-D rough-fault
motion in earthquake rupture zones. J. Geophys. Res. 103 (B11), 26885–26898,
Noda, H., Lapusta, N., 2013. Stable creeping fault segments can become destructive
the accretionary wedge. J. Geophys. Res. 97, 15321–15337.
Plafker, G., Savage, J.C., 2010. Comparisons of Near-Field Tsunami Characteristics of
Organized by Giant (M > 9) Earthquakes in Chile (1960), Alaska (1964), and Sumatra (2004).
Polet, J., Kanamori, H., 2000. Shallow subduction zone earthquakes and their
tsunamigenic potential. Geophys. J. Int. 142, 684–702, http://dx.doi.org/10.1046/j.1365-
Rivera, L., Kanamori, H., 2005. Representations of the radiated energy in earth-
Shi, Z., Day, S.M., 2013. Rupture dynamics and ground motion from 3-D rough-fault
Satake, K., Tanioka, Y., 1999. Sources of tsunami and tsunamigenic earthquakes in
Shi, Z., Day, S.M., 2013. Rupture dynamics and ground motion from 3-D rough-fault
Seno, T., 2000. The 21 September, 1999 Chi-Chi earthquake in Taiwan: implications
Shi, Z., Day, S.M., 2013. Rupture dynamics and ground motion from 3-D rough-fault
Kostrov, B.V., 1974. Seismic moment and energy of earthquakes, and seismic flow of
Lay, T., Bilek, S., 2007. Anomalous earthquake ruptures at shallow depths on
subduction zone megathrusts. In: Dixon, D.H., Moore, J.C. (Eds.). The Seismo-
genic Zone of Subduction Thrust Faults. Columbia University Press, New York,
NY, pp. 476–511.
122
Lay, T.H., Kanamori, C.A., Ammon, K.D., Koper, A.R., Hutko, L., Ye, H., Yue, Rushing, T.M.,
2012. Depth-varying rupture properties of subduction zone megathrust faults. J.
Ma, S., 2009. Distinct asymmetry in rupture-induced inelastic strain across dipping
earthquake: from real-time discriminants, natural science.280.5366.1055.
the accretionary wedge. J. Geophys. Res. 97, 15321–15337.
Plafker, G., Savage, J.C., 2010. Comparisons of Near-Field Tsunami Characteristics of
Organized by Giant (M > 9) Earthquakes in Chile (1960), Alaska (1964), and Sumatra (2004).
the accretionary wedge. J. Geophys. Res. 97, 15321–15337.
Plafker, G., Savage, J.C., 2010. Comparisons of Near-Field Tsunami Characteristics of
Organized by Giant (M > 9) Earthquakes in Chile (1960), Alaska (1964), and Sumatra (2004).
Poley, J., Kanamori, H., 2000. Shallow subduction zone earthquakes and their
tsunamigenic potential. Geophys. J. Int. 142, 684–702, http://dx.doi.org/10.1046/j.1365-
246X.2000.02095.x.
Rivera, L., Kanamori, H., 2005. Representations of the radiated energy in earth-
quakes. Geophys. J. Int. 162, 148–155, http://dx.doi.org/10.1111/j.1365-
246X.2005.02648.x.
Satake, K., Tanioka, Y., 1999. Sources of tsunami and tsunamigenic earthquakes in
Shi, Z., Day, S.M., 2013. Rupture dynamics and ground motion from 3-D rough-fault
Kostrov, B.V., 1974. Seismic moment and energy of earthquakes, and seismic flow of