The Initiation of Subduction: Criticality by Addition of Water?

Klaus Regenauer-Lieb,[1]* Dave A. Yuen,[2] Joy Branlund[3]

Subduction is a major process of plate tectonics; however, its initiation is not understood. We used high-resolution (less than 1 kilometer) finite-element models based on rheological data of the lithosphere to investigate the role played by water on initiating subduction. A solid-fluid thermomechanical instability is needed to drive a cold, stiff, and negatively buoyant lithosphere into the mantle. This instability can be triggered slowly by sedimentary loading over a time span of 100 million years. Our results indicate that subduction can proceed by a double feedback mechanism (thermoelastic and thermal-rheological) promoted by lubrication due to water.

The initiation of subduction remains one of the unresolved challenges of plate tectonics. Models of subduction (*1*, *2*) require an instability that forms a weak zone to initiate subduction. This first instability ruptures the mechanical coherence of the lithosphere by localized plastic yielding (*3*, *4*). Fluid-dynamical approaches have succeeded in producing near vertical drops of negatively buoyant viscous lithosphere promoted by such localized zones of weakness (*5*, *6*). However, the yield strength of the lithosphere has been assigned arbitrarily. The viscous-plastic analogy also neglects the importance of elastic bending, at oceanic trenches. Hence, an elasto-visco-plastic instability is required that overcomes both the barrier of flexural bending and shear resistance on the weak zone in a second finite-amplitude instability (*1*, *7*).

Here, we studied the nonlinear elasto-visco-plastic failure of the lithosphere from linear incremental loading of sediment piles at passive margins (*8*). Water content C_{OH} is the only parameter we vary in our finite-element model (*9*). The sediment load function (*10*) and the thermal model are fixed at the outset. Sedimentary rock and water buoyancy response forces are applied to the bottom of the plane strain model where two families (z and x) of linear dashpot elements (force–relative velocity couple) mimic an asthenosphere viscosity of 10^{21} Pa·s (Fig. 1).

The rheology of the oceanic lithosphere is a function of temperature, pressure, and water content. The constitutive equations for the lithosphere are expressed in tensorial form as (i) Peierls stress (*11*, *12*)

$$\dot{\varepsilon}_{ij}^{L} = A\alpha \exp\left[-\frac{Q_L + p(V_L - \frac{r}{k}\Delta V_{OH})}{RT}\left(1 - \frac{\tau_{ij}}{\tau_o}\right)^2\right] \quad (1)$$

(ii) power law creep (*13*)

$$\dot{\varepsilon}_{ij}^{P} = B\alpha J_2^{n-1}\tau_{ij} \exp\left[-\frac{Q_p + p(V_p - \frac{r}{k}\Delta V_{OH})}{RT}\right] \quad (2)$$

and (iii) diffusion creep (*14*)

$$\dot{\varepsilon}_{ij}^{D} = C\alpha J_2^{m-1} \frac{\tau_{ij}}{g^l} \exp\left[-\frac{Q_D + p(V_D - \frac{r}{k}\Delta V_{OH})}{RT}\right] \quad (3)$$

where A, B, and C are material constants. Values are given in Table 1. The weakening effect of water on rheology is twofold. It can be separated into a scalar weakening parameter α (*15*), responsible for increasing strain rate with increasing water content in the lithosphere, and its weakening for increasing pressure by the $\Delta V_{OH}-$ term (change in molar volume associated with the incorporation of hydroxyl ions into forsterite); r and k are fitting exponents introduced in (*13*, *14*, *16*). Q and V are activation energy and activation volume (where the subscripts refer to flow law), p is the magnitude of pressure defined by the first invariant (trace) of the stress tensor, T is the temperature, R is the universal gas constant, and g is the grain diameter of the crystal; n, m, and l are constants: the power-law creep, diffusion, and grain size exponents, respectively. τ_{ij} is the deviatoric stress and $\dot{\varepsilon}_{ij}$ is the corresponding strain rate tensor, which is coaxial to the flow potential defined by the von Mises equivalent stress, which is the second invariant J_2 of τ_{ij}. We assume that the pressure sensitivity of Peierls plasticity can be recast in terms of pressure sensitivity of the melting temperature of forsterite (*17*, *18*). Composite solid and fluid mechanical properties are implemented by a linear elastic spring in series with the fluid dynamic properties defined above. The total strain rate in plane strain becomes

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}_{ij}^{E} + \dot{\varepsilon}_{ij}^{T} + \dot{\varepsilon}_{ij}^{L} + \dot{\varepsilon}_{ij}^{P} + \dot{\varepsilon}_{ij}^{D} \text{ for } 2J_2 < p \quad (4)$$

The superscripts E and T refer to elastic and thermal expansion strain rates. At low temperatures, the creep stress overestimates the strength of the lithosphere and $2J_2 > p$. In this case, p defines the yield stress above which the composite fluid dynamic rheology is replaced by a von Mises elasto-plastic material (*11*).

The coupled temperature-displacement problem is fully solved with an energy equation:

$$\frac{DT}{Dt} = \kappa\nabla^2 T + \frac{\psi}{\rho C_p}\tau_{ij}\dot{\varepsilon}_{ij}^{F} \quad (5)$$

where DT/Dt refers to substantive derivative. The first term describes heat conduction, with κ the thermal diffusivity and ∇^2 the Laplace operator. The second term incorporates shear heating where ψ is the shear heating efficiency (*19*, *20*) of the dissipative processes in the fluid dynamic state $\dot{\varepsilon}_{ij}^{F} = \dot{\varepsilon}_{ij}^{L} + \dot{\varepsilon}_{ij}^{P} + \dot{\varepsilon}_{ij}^{D}$; above the creep threshold stress p, plastic strain rates replace the fluid dynamic term $\dot{\varepsilon}_{ij}^{F}$. C_p is the

Table 1. Parameter values as given in (*11*) through (*14*) without scaling for tensor notation. Peierls dry: $C = 5.7 \times 10^{11}$ s^{-1}, $\tau_0 = 8.5 \times 10^9$ Pa, $Q_L = 536$ kJ mol^{-1}. Peierls wet: $C = 1.0 \times 10^{12}$ s^{-1}, $\tau_0 = 9.1 \times 10^9$ Pa, $Q_L = 498$ kJ mol^{-1}.

Quantity	Diffusion	Power law
A, B	4.8×10^4 μm^3 MPa$^{-2.1}$ s^{-1}	1.5×10^3 μm^3 MPa^{-4} s^{-1}
V_D, V_P	20×10^{-6} m^3 mol	20×10^{-6} m^3 mol^{-1}
Q_D, Q_P	295 kJ mol^{-1}	470 kJ mol^{-1}
m, n	1.1	3
r	0.9	0.98
k	1 (0.9)	1 (0.9)

[1]Institute of Geophysics, Eidgenössische Technische Hochschule-Hönggerberg, Swiss Federal Institute of Technology, 8093 Zürich, Switzerland. [2]Minnesota Supercomputer Institute and Department of Geology and Geophysics, University of Minnesota, Minneapolis, MN 55415–1227, USA. [3]Southwestern Illinois College, 4950 Maryville Road, Granite City, IL 62040–2699, USA.

*To whom correspondence should be addressed. E-mail: klaus@tomo.ig.erdw.ethz.ch

specific heat at constant pressure and ρ is the density of forsterite.

Our formulation provides a self-consistent, laboratory data–based constitutive equation of a thermo-elasto-visco-plastic mechanical lithosphere (the solid layer) resting on a fluid dynamic rheological sublayer. The boundary between the fluid and solid domains is defined by the validity domain and the importance of the Peierls mechanism. In laboratory experiments, the boundary was found to be at about T_{base} = 1200 K corresponding to fluidlike deformation above and solidlike deformation below this temperature (*12*). Geological extrapolations expand the range to lower temperatures 1100 to 1200 K (*21*). In our formulation, the potential for solidlike faulting is embedded in the fluid layer through the yield phenomenon of the Peierls mechanism. Above the yield stress, the solid becomes weak and flow starts with a characteristic strain rate of

$$\dot{\varepsilon}_0^L = \alpha A \exp\left[-\frac{Q_L + p(V_L - \frac{r}{k}\Delta V_{OH})}{RT}\right] \quad (6)$$

We use a water content of $810 > C_{OH} > 80$ parts per million (ppm) H/Si corresponding to wet mid-ocean ridge basalt source material (*22*) and nominally dry material (*13*) to calculate the weakening parameter α (*15*) for power law and diffusion creep. The effect of water on Peierls plasticity is un-

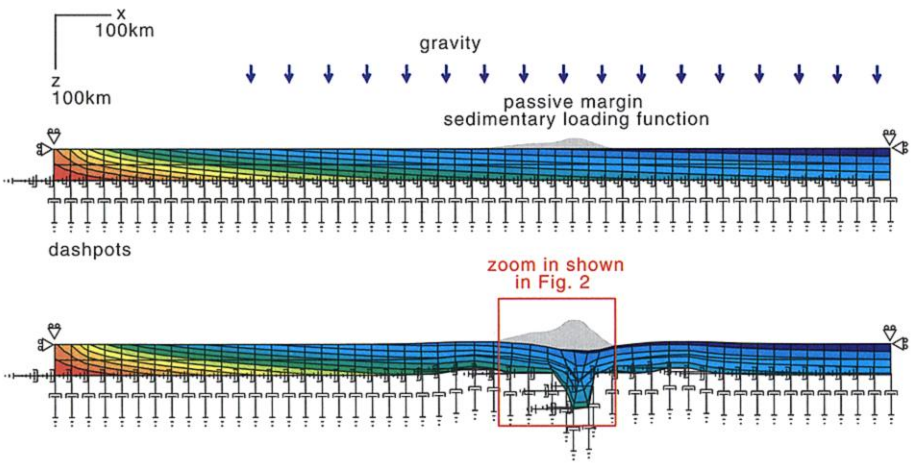

Fig. 1. Sketch of the model setup inspired by the Atlantic passive margin. The lithosphere is loaded by a steady flux of sediments; the peak load at 100 My is defined by summation of Gaussian functions $a_i \exp[-(x - s_i)^2/(2w_i^2)]$, with $a_{1,2}$ = 10 km, w_1 = 30 km, w_2 = 10 km, s_1 = 0, and s_2 = 30 km. The initial thermal conditions are given by the contoured cooling half-space model. The analysis focuses on weakening by rheological feedback mechanisms. Therefore, no geometrical/chemical seed for shear zone nucleation has been assumed other than sources of asymmetry by sediment load function and thermal profile. Deformation within the weak brittle top 10 km of the lithosphere, i.e., the sediment pile, is not considered; the top 10 km is stripped off.

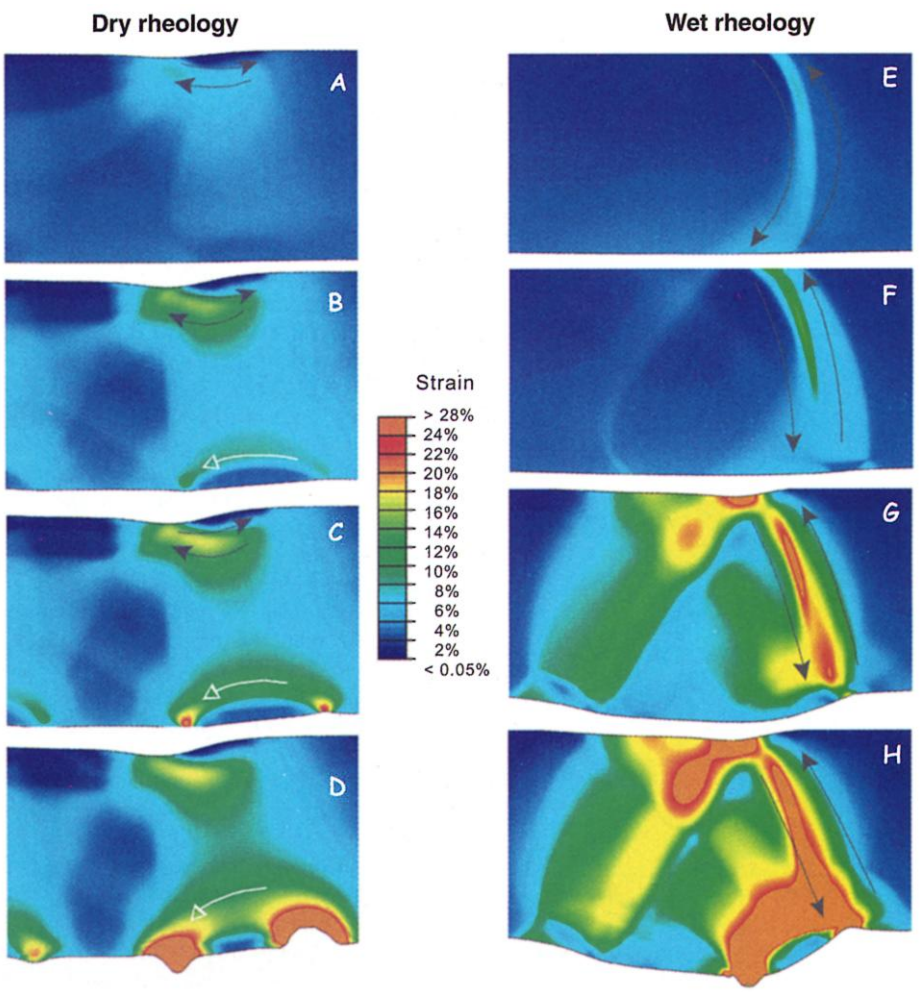

Fig. 2. Evolution of dry and wet oceanic lithosphere cases. Dry case (**A** to **D**) (C_{OH} = 80 ppm H/Si): (A) (38 My) shows a diffuse shear zone (paired arrows indicate shear sense) that nucleates near the model surface (10-km depth) after a 5.7-km-thick sediment pile has accrued. The shear zone decouples surface deformation from deep flow (white arrow) in (B) (64 My), (C) (69 My), and (D) (72 My). Rayleigh-Taylor instabilities near the base of the lithosphere finally lead to delamination of the bottom of the lithosphere, which drops off into the mantle. Wet case (**E** to **H**) (C_{OH} = 810 ppm H/Si): After 25 My of sediment loading (panel E), a hinge-like shear zone propagates through the lithosphere and ruptures its mechanical coherency. During ensuing highly localized deformation in (F) (30 My) and (G) (35 My), the hinge straightens up and rotates counterclockwise. In (H) (35.5 My), deformation along the hinge speeds up substantially. Strain weakening (not enabled here) ensures longevity of the shear zones indicated by saturated colors (orange > 28%). In (H), the shear zone reaches contours in excess of 250% strain within the saturated color range. The top of the lithosphere is weakened by a fault that zigzags in left and right lateral shear sense. During subsequent deformation, the left part plunges into the mantle, preserving asymmetry through elasticity, an important difference to fluid dynamic approaches (*5*, *6*).

known. We assume that (*11*) and (*12*) define the dry and wet rheology, respectively. With a linear scaling of α between these bounds and Eq. 6, we obtain a range of characteristic strain rates between $2 \times 10^{-16}\ s^{-1} < \dot{\varepsilon}_0^L < 3 \times 10^{-14}\ s^{-1}$ at $T_{base} = 1100$ K.

We investigate whether subduction initiation is possible for a reasonable tectonic load. Ridge push and sediment loads are possible contenders for breaking the two-dimensional geometry of plate margins. Ridge push results from thermal expansivity and acts as a far-field lateral push (line load of O 10^{12} N m^{-1}). Sediment loads on the contrary act locally and depress passive continental margins vertically. Erosion can reach an integrated load of the order of 10^{13} N m^{-1}; therefore, it should be more important than ridge push for bringing passive margins to criticality. To model the dynamics of the initial instability, we assumed that the continental margins are uniformly filled with sediments.

Both wet and dry lithosphere fail under an applied sediment load in our simulations. However the style of failure is different. In the dry lithosphere case (Fig. 2, A to D), only the top 10 km fails in a solid mechanical manner, whereas the lower fluidlike part, defined by $2J_2 < p$, deforms in a diffuse way. This does not favor subduction initiation. However, in the dry case, the sediment load triggers a Rayleigh-Taylor instability in the lower fluidlike portion of the lithosphere, the bottom of which delaminates after 70 million years (My) from the mechanical lithosphere and drops off. For the dry lithosphere case, our cooling half-space model thus develops dynamically into a cooling plate model between 70 and 100 My without precipitating subduction.

In the case of the wet lithosphere (Fig. 2, E to H), it fails on its entire mechanical thickness and is a suitable candidate for initiating subduction. The lower fluidlike part of the lithosphere can deform as a coupled entity together with the solidlike elasto-plastic upper part, whereas in the dry case it does not. The ability to maintain coherency in the deformation process is determined by the lithospheric strength and the role of the plastic Peierls stress mechanism within the viscous layer. In the dry case, an elastic core is present in the lithosphere. This prevents lithospheric separation for realistic sediment loads when considering additionally other possible sediment load functions and sedimentation-deformation feedback (*4*). The importance of the Peierls stress mechanism on triggering "dry" or "wet" modes of lithospheric failure is obvious from our parameter runs. Varying α within the two modes simply delays or accelerates the onset failure. However, the wet mode is switched on for all runs above a value of $\dot{\varepsilon}_0^L > 10^{-15}\ s^{-1}$ or $C_{OH} = 150$ ppm H/Si.

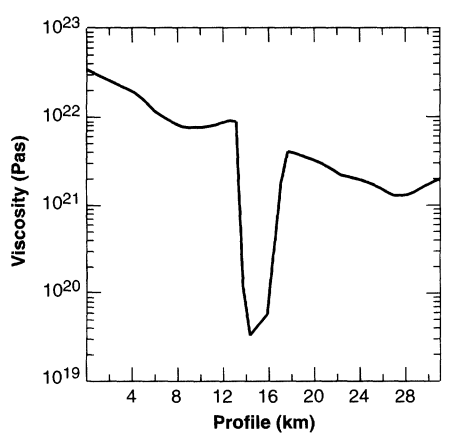

Fig. 3. Viscosity profile across the shear zone. The viscosity minimum marks its location.

Just by adding water, we obtain a narrow faultlike zone for lithosphere separation. It needs to be shown that the fault zone can be sufficiently softened to sustain subduction. The strain inside the major fault zone is large (Fig. 2, E to H) and well above the 5% suggested for strain-weakening mechanism such as void-volatile interaction (*23*, *24*) and dynamic recrystallization (*25*). Both mechanisms are promoted by the presence of water, but a sound quantitative description does not exist.

Therefore, we investigated only thermal-mechanical feedback due to the shear heating term in the energy equation (Eq. 5). Shear heating feeds back exponentially into the strain rate equations (Eqs. 1 to 3) and focuses high stress into the shear zone owing to thermal expansion (Eq. 4). Only a modest amount of shear heating is required (10 to 20 K) (*26*, *27*) to cause appreciable weakening. Even in the absence of a geometrical/chemical seed for shear zone localization (*28*), the addition of thermal-elastic expansion causes marked localization. Our present 600-m global resolution model can only give an upper bound of shear zone width and inferred strain rates ($10^{-12}\ s^{-1}$) or viscosity (10^{20} Pa·s), which is weaker than the model asthenosphere (10^{21} Pa·s) (Fig. 3).

We have shown that subduction initiation, and therefore, by inference, plate tectonics (*29*), rely on the presence of water. An Atlantic-type continental margin would reach two-dimensional criticality upon the addition of sediments, if an average load of 10-km thickness is reached. Plummeting of such a gravitationally and mechanically unstable lithosphere into the mantle can be envisioned through many subsequent three-dimensional mechanisms. Environmental loads (sediments, volcanic plateaus, and so forth) can cause Rayleigh-Taylor instabilities leading to removal of the bottom of the lithosphere for dry mantle cases. Positive feedback mechanisms hold the key for solving the subduction initiation paradox (*7*). We have shown that water and thermal-mechanical feedback can generate a narrow (<600 m) low-viscosity shear zone that cuts across the lithosphere.

References and Notes

1. J. Toth, M. Gurnis, *J. Geophys. Res.* **103**, 18053 (1998).
2. A. I. Shemenda, *J. Geophys. Res.* **98**, 16167 (1993).
3. J. Branlund, K. Regenauer-Lieb, D. Yuen, *Geophys. Res. Lett.* **27**, 1989 (2000).
4. ———, *Earth Planet. Sci. Lett.* **190**, 237 (2001).
5. L. Moresi, V. Solomatov, *Geophys. J. Int.* **133**, 669 (1998).
6. R. Trompert, U. Hansen, *Nature* **395**, 686 (1998).
7. D. P. McKenzie, in *Island Arcs Deep Sea Trenches and Back-Arc Basins*, M. Talwani and W. C. Pitman, Eds., vol. 1 of Maurice Ewing Series (American Geophysical Union, Washington, DC, 1977), pp. 57–61.
8. S. A. P. L. Cloetingh, M. J. R. Wortel, N. J. Vlaar, *Nature* **297**, 139 (1982).
9. ABAQUS/Standard, User's Manual, vol. 1, version 6.1. (Hibbit, Karlsson and Sorenson, Pawtucket, RI, 2000).
10. B. E. Tucholke, R. E. Houtz, W. J. Ludwig, *AAPG Bull.* **66**, 1393 (1982).
11. C. Goetze, B. Evans, *Geophys. J. R. Astron. Soc.* **59**, 463 (1979).
12. B. Evans, C. Goetze, *J. Geophys. Res.* **84**, 5505 (1979).
13. S. Mei, D. L. Kohlstedt, *J. Geophys. Res.* **105**, 21471 (2000).
14. ———, *J. Geophys. Res.* **105**, 21457 (2000).
15. We use the molar concentration of hydroxyl C_{OH} to define $\alpha^r = (C_{OH}/1.1)^k$ and $V_{OH} = 10.6 \times 10^{-6}$ m^3 mol^{-1} for $r = 1$.
16. D. L. Kohlstedt, H. Kepler, D. C. Rubie, *Contrib. Mineral. Petrol.* **123**, 345 (1996).
17. S. Karato, D. Yuen, M. R. Riedel, *Phys. Earth Planet. Inter.*, in press.
18. D. C. Presnall, M. Walter, *J. Geophys. Res.* **98**, 19777 (1993).
19. K. Regenauer-Lieb, J. P. Petit, D. Yuen, *Electron. Geosci.* **4**, 2 (1999).
20. The strain-energy rate density (double product of the strain rate and stress tensor) can be decomposed into dissipated energy, appearing as heat, and a complementary part that is stored during deformation. The shear heating efficiency is about 90% of the total strain-energy rate density. In our calculations, we only consider the dissipative part and neglect the stored energy.
21. G. Ranalli, *Tectonophysics* **240**, 107 (1994).
22. G. Hirth, D. L. Kohlstedt, *Earth Planet. Sci. Lett.* **144**, 93 (1996).
23. K. Regenauer-Lieb, *J. Geodyn.* **27**, 1 (1999).
24. D. Bercovici, *Earth Planet. Sci. Lett.* **154**, 139 (1998).
25. J. Braun *et al.*, *J. Geophys. Res.* **104**, 25167 (1999).
26. D. C. Roberts, D. L. Turcotte, in *GeoComplexity and the Physics of Earthquakes*, J. B. Rundle, D. L. Turcotte, W. Klein, Eds. (American Geophysical Union, Washington, DC, 2000), pp. 97–104.
27. K. Regenauer-Lieb, D. Yuen, *Phys. Earth Planet. Inter.* **118**, 89 (2000).
28. ———, *Geophys. Res. Lett.* **25**, 2737 (1998).
29. W. M. Kaula, *Philos. Trans. R. Soc. London Ser. A Math. Phys. Eng. Sci.* **349**, 345 (1994).

30. We thank S.-I. Karato, D. L. Kohlstedt, S. Goes, G. Morra, and F. Funiciello for helpful discussions. This is contribution 1194 of the Institute of Geophysics Eidgenössische Technische Hochschule Zurich.

28 June 2001; accepted 17 September 2001