Studies in Fluid Dynamics as Applied to Seismology and Volcanology

Thesis by

Emily E. Brodsky

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2001
(Submitted October 25, 2000)

This thesis is dedicated to Bradford Sturtevant (1933–2000).

Acknowledgments.

My advisors, Hiroo Kanamori and Brad Sturtevant, have provided superb advice both jointly and individually. It is in a large part due to their insight, encouragement and cooperation that the past five years have been so productive.

During my graduate career I have been infected by the intense curiosity fostered by the culture of the Caltech Seismo Lab. I thank everyone who has participated in coffee discussions for sharing both their criticism and enthusiasm.

The core of this thesis is a series of submitted and published papers that have been expanded and clarified. The co-authors of these papers were an integral part of the work. They are: Hiroo Kanamori (chapters 2–5), Brad Sturtevant (chapters 4–5) and Vassilis Karakostas (chapter 3). I would also like to acknowledge the following people for the contributions to individual chapters:

Chapter 2: Reviews by Y. Ben-Zion, J. Brune, V. Lyakhovsky and N. Sleep were insightful and constructive.

Chapter 3: An introduction by L. Rivera made this collaboration possible.

Chapter 4: Comments by T. Koyaguchi, S. Kaneshima, M. Rutherford and J. Lowenstern greatly improved an early version of this manuscript. More recent insights by M. Manga and especially M. Ichihara clarified the limitations of the theory.

Chapter 5: Correspondence and conversations with S. Kieffer, S. Malone, R. Waitt and R. Hoblitt were valuable in understanding the geological and observational records.

I also appreciate the personal support that my family, officemates and friends provided through some particularly difficult times. The reading group especially has my gratitude for providing an oasis of literature and gossip. Francis Nimmo deserves acknowledgment for his many contributions to this thesis which include proofreading it in entirety, feeding its author, cleaning its author's dishes and caring for its author's cats. I am, as always, grateful.

Abstract

This thesis addresses three problems in seismology and volcanology by applying fluid dynamical theories that have been developed for engineering applications.

Fault zones are proposed to operate analogously to journal bearings. A quantitative assessment of the physical regimes in which faults behave as lubricated systems is made using elastohydrodynamic theory. Elastohydrodynamic lubrication with typical parameters explains the following observable phenomenon: (1) a reduction in the frictional stress by 50% during large earthquakes, (2) a decrease in high-frequency (>1 Hz) radiation above a critical slip distance of a few meters and (3) a two orders of magnitude variation in scaled radiated energy between small ($M_w < 4$) and large earthquakes ($M_w > 6$).

Regionally triggered seismicity often occurs in geothermal areas. It is documented here that the 1999 M_w=7.4 Izmit, Turkey, earthquake was followed by widespread seismicity in Greece over a study region extending from 400 km to nearly 1000 km away from the epicenter. The increase in cataloged earthquakes is statistically significant at the 95% level. A related phenomenon is the regional triggering of volcanic eruptions. A model for triggering eruptions based on rectified diffusion is formulated and evaluated. The excess pressure from rectified diffusion in a typical basaltic system following a regional $M \geq 8$ earthquake is between 0.001 and 0.02 MPa. Strong constraints on the porosity, size of the bubbly region, velocity structure and permeability must be imposed for rectified diffusion to be effective.

A fluid dynamical model based on supersonic nozzle flow is used to link observed seismic waves with the mass discharge rate of an explosive volcanic eruption. The method is tested by calculating the vertical mass discharge rate from Mount St. Helens for the beginning of the May 18, 1980 eruption. The observed seismic sources are modeled as thrusts due to a combination of the momentum flux of the erupted products and the pressure of the eruptive jet. The momentum discharge rate is con-

verted to a mass discharge rate. The calculated mass ejected in the first 100 s is 1.6×10^{11}–4.6×10^{11} kg. Since the total blast deposit is $\sim3.2\times10^{11}$–4.1×10^{11} kg, one interpretation is that the directed blast had a significant ($\geq$40%) vertical component.

Contents

List of Figures

Chapter 1 Overview

This thesis aims to demonstrate that fluid dynamics is an important component of solid Earth geophysics. The contributions of fluid dynamics have long been appreciated in certain subdisciplines, such as the study of the mantle. Over the last 20 years, fluid dynamical studies have become a significant component of modern volcanology [e.g., *Wilson et al.*, 1978; *Sparks et al.*, 1997]. Seismology has remained virtually untouched. In this thesis I show that engineering studies of fluid mechanics conducted for industrial applications have important ramifications for both the qualitative and quantitative study of seismology and volcanology.

Part I discusses the elastohydrodynamic lubrication of fault zones. Lubrication may be an important effect in reducing the normal stress during rupture and therefore the process helps to explain the apparently low friction on faults. This line of research is an example of applying engineering methods and data to a geophysical system. Lubricated bearings have been intensively studied since the seminal work by Reynolds [*Reynolds*, 1886] and the parallels between the artificial and natural systems are striking. Chapter 2 includes quantitative assessments of the physical regimes in which faults behave as lubricated systems by using elastohydrodynamic theory and the nondimensional Sommerfeld number [*Sommerfeld*, 1950]. The theory provides a unified explanation for a number of disparate seismological observables such as the variations in radiated energy between large and small earthquakes and the nearfield waveforms of large earthquakes such as the 1999 Chi-Chi, Taiwan, $M_w = 7.6$ event.

Recent observations have indicated that seismicity is commonly triggered hundreds of kilometers from a mainshock epicenter. This puzzling phenomenon is not well explained by traditional elastic models of seismic stresses. The fact that the sites of triggered seismicity are commonly associated with geothermal and magmatic activity suggests that fluid dynamics may be an important component of the triggering processes. Part II includes both an observational and theoretical study of such regional triggering. Chapter 3 documents that the 1999 Izmit earthquake was followed by widespread seismicity up to 1000 km away from the epicenter and that the triggered events often occurred in geothermal areas. Similar correlations between triggered seismicity and geothermal regions were observed after the 1992 Landers and 1999 Hector Mine earthquakes. Chapter 4 evaluates a model to explain triggering by seismic waves based on the interactions of waves with bubbly fluids. The particular mechanism explored here is known as "rectified diffusion." It has been studied intensively in fields such as bioengineering [*Leighton*, 1994], but the basic theory required modification for use in magmatic systems. The new formalism both extends the original fluid dynamical studies and contributes to our understanding of the interplay between processes in magmatic systems.

Part III develops a method for using seismic data to invert for the mass ejection rate of explosive volcanic eruptions. The momentum discharge rate of a volcanic eruption is balanced by a force into the ground which generates seismic waves. Instrumental records of the waves can therefore be used to invert for this momentum discharge rate. A simple fluid dynamical model based on supersonic flow through a nozzle is used to convert the momentum to mass discharge. This fluid dynamical model has long been used for volcanic systems [*Kieffer*, 1981], but a direct link between the data and the fluid dynamic model is a new contribution. The resulting mass ejection rate in kilograms per second is a fundamental and useful metric of explosive eruptions. This method, which I call thrust inversion, combines traditional seismological methods with fluid dynamics to formulate the first entirely instrumental metric of explosivity.

The projects in this thesis represent some exciting new directions that are motivated by recent observations. Insights into the fundamental physics of all of these phenomena are gained by approaching the fluid dynamics of the solid Earth with the tools developed by engineering science.

References

Kieffer, S. W., Fluid dynamics of the May 18 blast at Mount St. Helens, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 379–400, 1981.

Leighton, T., *The Acoustic Bubble*, Academic, San Diego, Calif., 1994.

Reynolds, O., On the theory of lubrication and its application to Mr. Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil, *Phil. Trans. Roy. Soc.*, *177*, 157–234, 1886.

Sommerfeld, A., *Mechanics of Deformable Bodies*, Academic Press, New York, 1950.

Sparks, R. S. J., M. I. Bursik, S. N. Carey, J. Gilbert, L. Glaze, H. Sigurdsson, and A. W. Woods, *Volcanic Plumes*, John Wiley, New York, 1997.

Wilson, L., R. S. J. Sparks, T. C. Huang, and N. D. Watkins, The control of volcanic column heights by eruption energetics and dynamics, *J. Geophys. Res.*, *83*, 1829–1836, 1978.

Part I

Lubrication of Faults

Chapter 2 The Elastohydrodynamic Lubrication of Faults

Abstract

The heat flow paradox provides evidence that a dynamic weakening mechanism is important in understanding fault friction. We present here a specific model for dynamic velocity weakening that uses the mechanics of well-studied industrial bearings to explain fault zone processes. As two-subparallel solid surfaces shear a viscous fluid, an elevated fluid pressure is generated in the gap. This lubrication pressure supports part of the load, therefore reducing the normal stress and associated friction across the gap. The pressure also elastically deforms the wall rock. The model is parameterized using the Sommerfeld number, which is a measure of the lubrication pressure normalized by the lithostatic load. For typical values of the material properties, slip distance and velocity, the Sommerfeld number suggests that lubrication is an important process and reduces the frictional stress by 50% during large earthquakes. Elastohydrodynamic lubrication also predicts a decrease in high-frequency (>1 Hz) radiation above a critical slip distance of a few meters. This prediction is well-matched by the strong motion data from the 1999 Taiwan earthquake. The observed two orders of magnitude variation in scaled radiated energy between small ($M_w < 4$) and large earthquakes ($M_w > 6$) is also predicted by the lubrication model.

Introduction

It is a long-standing problem in seismology that frictional stresses on faults appear to be lower than expected based on laboratory evidence. Experiments measure the frictional stress of dry rocks sliding against rocks to be approximately 60% of the normal stress whereas heat flow data taken near the San Andreas fault suggest substantially less friction in natural systems [*Lachenbruch and Sass*, 1980]. High pore fluid pressures are often proposed to resolve this dilemma. Interstitial fluids partially support the load between the fault planes and therefore reduce frictional resistance. Models generally use the static values of pore pressure for this purpose. Here we consider the dynamic effect of fluids during seismic faulting. It is commonly observed in engineering applications that fluid pressure increases during motion [*Reynolds*, 1886]. As two sub-parallel solid surfaces slide past each other, they strain the viscous fluid between them. The finite viscosity of the fluid resists the motion and a high pressure gradient is formed. This pressure produces a normal stress that can help support the load. This hydrodynamic lubrication works in two ways to reduce the total frictional stress on the fault. (1) In the fraction of the fault area lubricated by a continuous film, the shear stress is the viscous stress which is significantly below the solid-solid friction for fluids such as water, slurry or mafic melt. (2) More importantly, the mechanically increased pore pressure reduces the effective stress on the solid-solid interfaces and therefore dynamically reduces the frictional stress during an earthquake.

In this study we develop a theory to quantitatively predict the frictional effects of coseismic lubrication by a viscous fluid in a confined fault. We then apply the theory to three specific observations: heat flow, nearfield waveform complexity and radiated energy. The fault is envisioned as a bicontinuous system of solid matrix and fault fluid. The fault fluid flows around asperities in a nearly linear path as the area of the asperity contacts is much less that the total fault surface area (Figure 2.1). The flow is modeled as two-dimensional and the deflection due to the asperities is neglected. We calculate the fluid pressure generated by this flow and then consider its effect on reducing the load on the asperities.

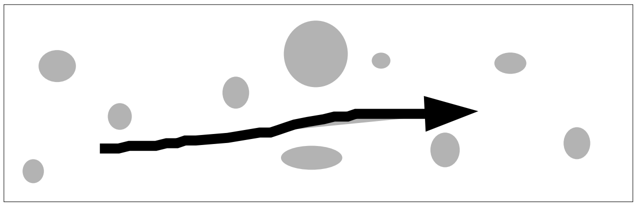

Figure 2.1 Map view of a fault surface. There is a bicontinuous system of solid matrix (rock) and fault fluid (slurry). The upper block (not shown) moves to the right relative to the lower block. The fluid is dragged around the asperities (grey) in the nearly slip-parallel path direction shown by the arrow.

The fault fluid is formulated as generally as possible since the physical process described refers to any viscous fluid sheared between two rough surfaces. We favor identifying the fluid as a slurry of fine-grained fault gouge and aqueous fluid because geological evidence for such a material exists [e.g., *Otsuki*, 1999] and slurry viscosities are large enough to produce significant lubrication pressures. Other candidate fault fluids include aqueous fluid and frictional melt. The general theory presented here applies equally well to any viscous fluid, although the numerical results would need to be recomputed with the appropriate material properties.

Lubrication theory

A quantitative understanding of hydrodynamic lubrication requires a brief review of the standard formulation of the motion of the fluid between two subparallel planes. The fluid motion is completely described by the full Navier-Stokes equation,

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla p + \eta \nabla^2 \mathbf{u}, \tag{2.1}$$

where ρ is the density, $\mathbf{u}$ is the velocity vector, p is the pressure and η is the viscosity. An important simplification of the Navier-Stokes equations can be made if the flow

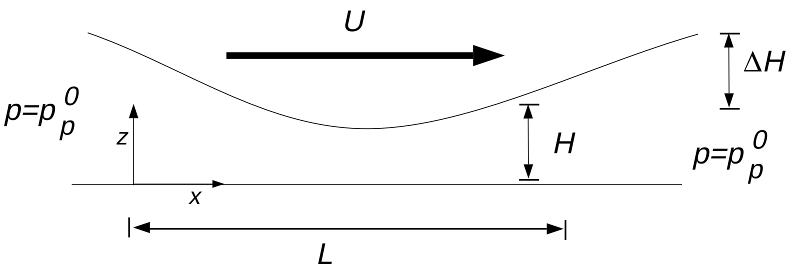

Figure 2.2 Characteristic dimensions H and L in a narrow fault. L is the characteristic dimension of pressure changes and H is the mean height of the gap. For a fluid-filled fault, the mean gap height is the same as the average thickness of the slipping zone. U is the slip velocity of one side of the fault relative to the other, p_p^0 is the undisturbed fluid pressure and ΔH is the average asperity height. All cartoons in this chapter are vertically exaggerated.

is through a thin gap, such as a fault zone. A thin gap is defined as one where the length, L, over which the fluid pressure changes significantly is much longer than the mean height, H, of the gap (Figure 2.2). The thin gap simplification of the equations of motion is known as the lubrication approximation and all terms in the Navier-Stokes equation of order H^2/L^2 or smaller are neglected. The inertial term is also negligible provided that the Reynolds number is much less than L^2/H^2 (Appendix A). For faults during rupture both conditions are met and the governing equations in two dimensions reduce to

$$\frac{\partial p}{\partial x} = \eta \frac{\partial^2 u}{\partial z^2} \tag{2.2}$$

$$\frac{\partial p}{\partial z} = 0 \tag{2.3}$$

where u is the fluid velocity in the x direction and the coordinate system adopted defines x as parallel to the fault slip vector and z as the normal to the fault plane (Figure 2.2).

Pressure gradients across the gap are negligible under the above assumptions and the dynamics are dominated by the balance of the viscous stresses and the pressure gradient parallel to the gap. The continuity equation for an incompressible fluid is

$$\frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}=0 \tag{2.4}$$

where w is the velocity in the z direction. Combining (2.4) with the boundary conditions and the equations of motion (2.2) and (2.3) provides solutions for the velocity and pressure fields.

We adopt a frame of reference that is stationary at the wall where $z = 0$ and assume no slip boundary conditions at the walls.

$$\begin{aligned}&\text{At } z=0, u=0, w=0\\&\text{At } z=h, u=U, w=0\end{aligned} \tag{2.5}$$

At the wall with $z = h$, where h is defined as the local slipping zone thickness, $u = U$ where U is the relative velocity between the fault walls. The relative motion between the walls is assumed to be entirely in the plane of the fault. The slipping zone height includes both the initial distance between the walls and any elastic deformation that may be caused by the fluid pressurization. At both walls ($z = 0$ and $z = h$), $w = 0$ since there is no flow in the direction normal to the impermeable walls.

We integrate (2.2), impose the boundary conditions (2.5), and use the integrated form of the continuity equation (2.4) to derive the Reynolds equation, a standard result in lubrication theory [e.g., *Hamrock*, 1994],

$$\frac{d}{dx}\left(h^3\frac{dp}{dx}\right)=-6\eta U\frac{dh}{dx}. \tag{2.6}$$

A more convenient integrated form is

$$p(x)-p_p^0=6\eta U\int_0^x\frac{h^*-h}{h^3}dx \tag{2.7}$$

where h^* is the gap height at the point of maximum pressure, i.e., at $x = x_*$ where $dp/dx|_{x=x_*} = 0$. Given boundary conditions for p and the geometry of the fault

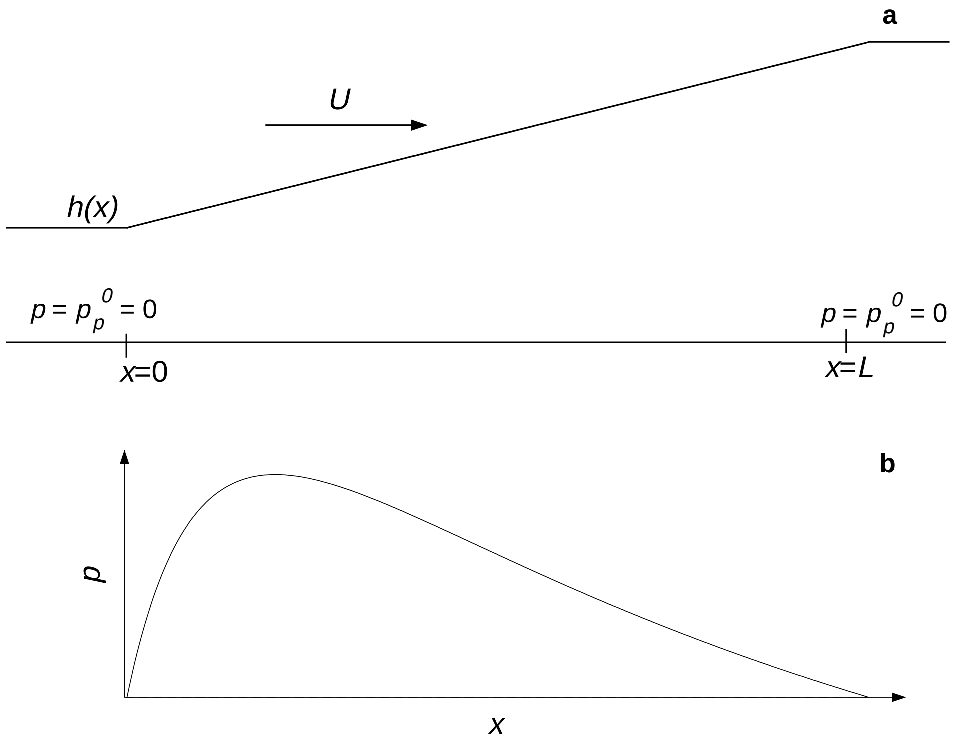

Figure 2.3 Lubrication under a slider block. (a) The geometry of the slider plate. (b) The pressure increase as calculated from (2.7) in arbitrary units.

specified by $h(x)$, the Reynolds equation can be solved to find the pressure distribution $p(x)$. The typical behavior of lubricated systems can be illustrated by considering a tapered slider moving over a plane surface (Figure 2.3).

The pressure in the farfield is the initial reservoir pressure, p_p^0. It is convenient to reference all pressures to this level and set p_p^0 equal to 0. The appropriate boundary conditions for pressure are that at $x = 0$ and $x = L$, the pressure $p = p_p^0 = 0$. The pressure rises in the narrow region and therefore exerts a net force or "lift" separating the two blocks.

Elastic effects

To this point we have assumed that the fault wall behaves as a rigid body. This is not true in many lubrication problems. The fluid pressure deforms the wall and in so doing adjusts the geometry to make lubrication more effective. This phenomenon explains how originally symmetric asperities like that pictured in Figure 2.2 deform to generate a net pressure increase. The everyday experience that arbitrary rough surfaces slide more easily when lubricated than when dry is evidence of the efficacy of the elastic adjustment.

The elastic deformation is computed by considering the fluid pressure to be a series of pressure line sources. The pressurization lasts as long as the fault is rupturing. Since the asperities are at most a few meters long, the time it takes the elastic waves to traverse the asperities ($\ll 10^{-2}$s) is likely much less than the duration of loading. Therefore, a quasistatic solution to the elastic equations is appropriate. Elastodynamic deformation is simultaneously occurring on the fault walls during rupture due to additional processes such as asperity collisions. Those additional dynamic effects are beyond the scope of the present study and would simply be superposed on the linear elastostatic solution presented here. The elastostatic solution is derived in Appendix B, and at point x the gap changes in height by a quantity $\delta(x)$,

$$\delta(x) = \frac{4(1-\nu^2)}{\pi E} \int_0^L p(\xi) \ln \left| \frac{a + \sqrt{(x-\xi)^2 + a^2}}{x - \xi} \right| d\xi, \tag{2.8}$$

where $2a$ is the width of the lubricated zone in the fault plane direction normal to slip, E is the Young's modulus of the rock and ν is the Poisson's ratio. We assume that the lubricated zone is equidimensional. The width is comparable to the length and $a = L/2$. We also assume that the pre-earthquake pore pressure p_p^0 is in equilibrium with the initial asperity shape and produces no deformation.

Figure 2.4 shows the deformation of an asperity with the initial shape $s(x) = h_1(\cos(2\pi x/L) + 1) + h_0$ where the constants h_1 and h_0 are both 10^{-3} m for this example. Details of the numerical method are in Appendix C. There are two features

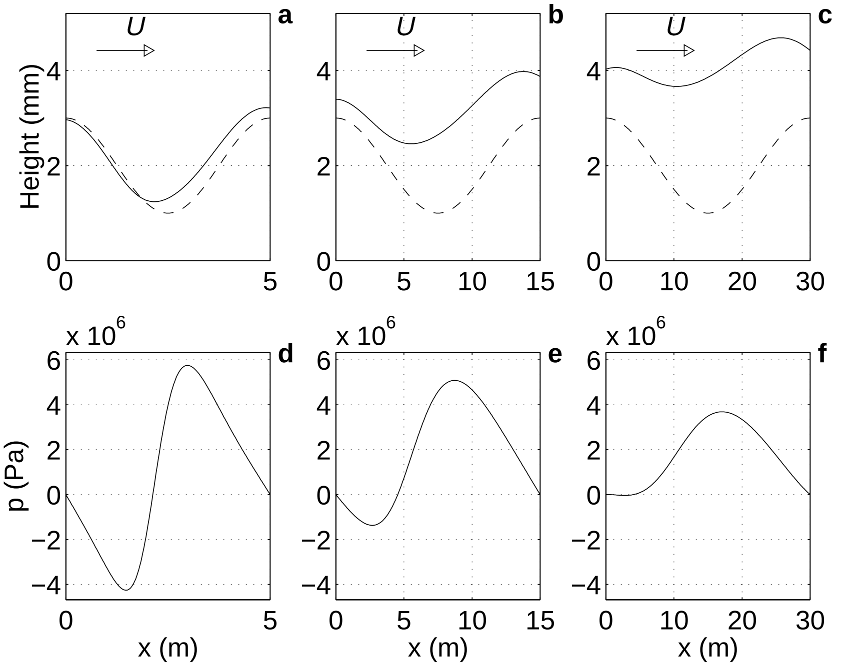

Figure 2.4 Elastic deformation of a cosine asperity with an initial minimum separation of 1 mm. (a) The dashed line is the undeformed asperity shape $s(x)$ and the solid line is the equilibrium shape $h(x) = s(x)+\delta(x)$ with the elastic deformation coupled to the lubrication fluid pressure. Parameters are $E = 5 \times 10^{10}$ Pa, $L = 5$ m, $U = 1$ m/s and $\eta = 10$ Pa s. (b) The same as (a) with $L = 15$ m. (c) The same as (a) with $L = 30$ m. (d) The change in pressure in the fluid due to lubrication effects. The total change in gap height is plotted and includes the elastic deformation of both sides of the faults. (e) Same as (d) with $L = 15$ m. (f) Same as (d) with $L = 30$ m.

of the solution that are particularly important for the discussion that follows. (1) The asperity is deformed to an asymmetrical configuration, thus promoting lubrication and supporting a net load. (2) As the length of the asperity L increases, the displacement $\delta(x)$ increases in accordance with (2.8). Since the initial asperity height is the same in all three cases of Figure 2.4, the deformed asperity flattens with increasing values of L.

Sommerfeld number

Equation (2.6) balances the pressure gradient in the gap with the viscous stresses generated by the variations in gap height and implies that the excess pressure due to lubrication, $P_L \equiv \mathcal{O}(p(x) - p_p^0)$, scales as

$$P_L \sim \frac{6\eta U L \Delta H}{H^3} \tag{2.9}$$

where H is the mean gap height and $\Delta H \equiv \mathcal{O}(h - h^*)$ is the mean variation in gap height caused by asperities (Figure 2.2). Capital letters are used here to denote characteristic scales while variables are written in lowercase. The lubrication pressure increases with viscosity. This dependence on viscosity is why oil is a better lubricant than water for most household uses [*Persson*, 1998, p. 97].

For a fault with asperities that intermittently are in contact during rupture, the average asperity height, ΔH, may be of the same order as the average gap height, H. The expression for the lubrication pressure would reduce to

$$P_L \sim \frac{6\eta U L}{H^2}. \tag{2.10}$$

However, the assumption that the mean asperity height ΔH is of the same order as the initial gap height H_0 is always not valid for natural surfaces. *Power and Tullis* [1991] showed that faults are self-similar over a wide range of scales. According to their measurements, $\Delta H = KL$ for values of L between 10 μm and 40 m where K is a constant of order 10^{-4}–10^{-2}. In this case, equation (2.9) reduces to

$$P_L \sim \frac{6\eta UKL^2}{H^3}. \tag{2.11}$$

For natural fault surfaces and large ruptures (earthquake magnitude, M, > 6), the parameter K is of the same order as H/L. As will be discussed below, lubrication pressure is only significant for large earthquakes. Therefore, equation (2.10) is often a convenient approximation for the lubrication pressure, even though equation (2.11) is more rigorously correct.

The effectiveness of lubrication is estimated by normalizing (2.11) by the pressure from the static load to form the dimensionless Sommerfeld number,

$$S \equiv \frac{P_L}{P} = \frac{6\eta UL^2K}{PH^3} \tag{2.12}$$

where P is the lithostatic pressure for geological problems [*Sommerfeld*, 1950]. This form of the Sommerfeld number is slightly modified from the original reference because of the geometric difference between the axisymmetric journal bearing problem that Sommerfeld was considering and the natural, planar fault considered here.

The average gap height H is equal to the sum of the initial average gap height H_0 and the average elastic deformation, D. According to Hooke's law, D is related to the lubrication pressure by

$$D = LP_L/E \tag{2.13}$$

where E is Young's modulus. For small lubrication pressures, the elastic deformation is negligible and

$$S = \frac{6\eta UKL^2}{(H_0)^3}\frac{1}{P}. \tag{2.14}$$

The initial gap geometry drives the lubricating flow and the asperities are undeformed. This regime occurs when the lubrication pressure is much less than the pressure necessary to generate deformation of order H_0,

$$P_L \ll EH_0/L. \tag{2.15}$$

We define a critical lubrication length L_c below which the elastic deformation D is insignificant. L_c is defined as the length at which $D = H_0$. Combining equations (2.11) and (2.13),

$$L_c = 2H_0 \left(\frac{H_0 E}{6\eta U K} \right)^{1/3}. \tag{2.16}$$

In the inelastic regime where $L \ll L_c$, the larger the value of L, the larger the lubrication pressure.

For large $L \gg L_c$, elastic deformation is significant. The initial gap height H_0 is much less than D and the total gap height H is primarily determined by the elastic contribution, i.e., $H \approx D$. The roughness K is assumed constant and therefore for $L \gg L_c$

$$S = \left(\frac{6\eta U K E^3}{L} \right)^{1/4} \frac{1}{P}. \tag{2.17}$$

Equation (2.17) indicates that S decreases with increasing L in the elastic regime. As the value of L increases, the elastic deformation widens the gap and therefore the lubrication pressure gradually decreases. This decrease of lubrication pressure in the elastic regime is a gradual process ($P_L \sim L^{-1/4}$) since the decreasing lubrication pressure also reduces the strain, D/L.

The Sommerfeld number measures the importance of lubrication in determining the frictional properties of a system. The effect of lubrication pressure on friction can be understood in terms of the classical model of *Hubbert and Rubey* [1959] describing the effects of pore fluid pressure. The effective pressure P_e on the solid contacts is the static load less the average pore pressure P_p [*Hubbert and Rubey*, 1959],

$$P_e = P - P_p. \tag{2.18}$$

The *Hubbert and Rubey* [1959] model, like nearly all frictional models, is based on the observation that contacting asperities account for the majority of the frictional stress even though they comprise only a small fraction of the fault plane area. This work shows that during rupture the pore pressure is dynamically increased and $P_p = P_p^0 + P_L$. If the Sommerfeld number is high, then P_L is significant and the effective pressure on contacting asperities is decreased. The contribution of the asperity contacts to the average frictional stress τ_f on the fault is

$$\tau_f = \mu_s P_e \tag{2.19}$$

where μ_s is the solid frictional coefficient.

For small Sommerfeld numbers, the lubrication pressure supports an insignificant fraction of the load. The two sides of the fault are in contact at asperities and conventional formulations of solid friction as proportional to the normal stress are appropriate. This limiting behavior is known as boundary layer friction. The frictional stress τ_f is determined by adhesive forces and can to some extent be predicted from the surface chemistry [*Persson*, 1998]. Boundary layer friction is conventionally formulated in terms of the frictional coefficient,

$$\tau_f = \mu_s P, \tag{2.20}$$

where μ_s is a material property of the solid.

For large Sommerfeld numbers, the lubrication pressure completely supports the load and no normal stress is exerted on contacting asperities. The only frictional stress is the viscous resistance of the fluid and (2.19) is not relevant. The magnitude of the viscous stress in the fluid is $\eta du/dz$. Equation (2.2), which is the lubrication equation of motion in the x direction, shows that the velocity gradient du/dz scales as $P_L H/L\eta$. Therefore, the viscous stress scales as $P_L H/L$. Since this viscous stress is the only frictional stress in the fully lubricated regime, the effective frictional coefficient $\mu \equiv \tau_f/P$ is

$$\mu = SH/L \tag{2.21}$$

where the definition of the Sommerfeld number as $S = P_L/P$ has been used to simplify the expression.

Between the boundary layer and hydrodynamic regimes is the “mixed regime.” Friction is determined by a mixture of viscous resistance and solid asperity contacts

(Figure 2.1). The effects are additive and the frictional stress is

$$\tau_f = \mu_s P_e + SPH/L. \tag{2.22}$$

If we neglect the initial hydrostatic fluid pressure P_p^0 for the purposes of this qualitative discussion of friction, then $P_e \approx P(1-S)$ and the effective frictional coefficient is

$$\mu = \mu_s(1-S) + SH/L. \tag{2.23}$$

All three types of behavior are shown schematically in the classical Stribeck curve in Figure 2.5 [*Spikes*, 1997] and are summarized by

$$\mu = \begin{cases} \mu_s & S \ll 1 \\ \mu_s(1-S) + SH/L & S \approx 1 \\ SH/L & S \gg 1. \end{cases} \tag{2.24}$$

The Stribeck curve demonstrates both effects of lubrication outlined in the introduction. In the hydrodynamic regime, the shear stress is low since solid-solid friction has been eliminated in favor of viscous effects. In the mixed regime, the friction has been depressed since the lubrication pressure is high.

Rough surfaces

The length of the lubricated zone L is defined as the length over which the pressure returns to its reservoir value. For a single asperity, such as shown in Figure 2.4, this is clearly the asperity length. Real fault surfaces are thought to be fractals with asperities over a continuous range of lengths [*Scholz*, 1990; *Power and Tullis*, 1991]. The ratio of asperity height to wavelength is constant to a first approximation [*Power and Tullis*, 1991]. Figure 2.6a is a simulation of a rough surface generated by the function

$$f(x) = \sum_{n=1}^{n_{\max}} \frac{K\lambda}{n} \cos(\frac{2\pi n x}{\lambda} + \phi_n) \tag{2.25}$$

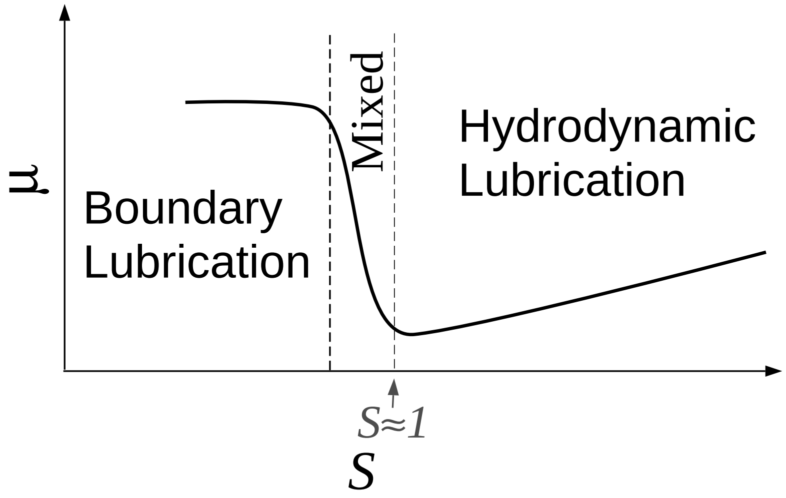

Figure 2.5 The Stribeck curve schematically shows the variation of the coefficient of friction, μ, with S. In the region labeled "boundary lubrication," solid-solid friction predominates. Between the dashed lines is the mixed regime where solid-solid friction is reduced by the lubrication pressure. In the "hydrodynamic lubrication" region, solid-solid friction is eliminated and the friction is determined by the viscous stress.

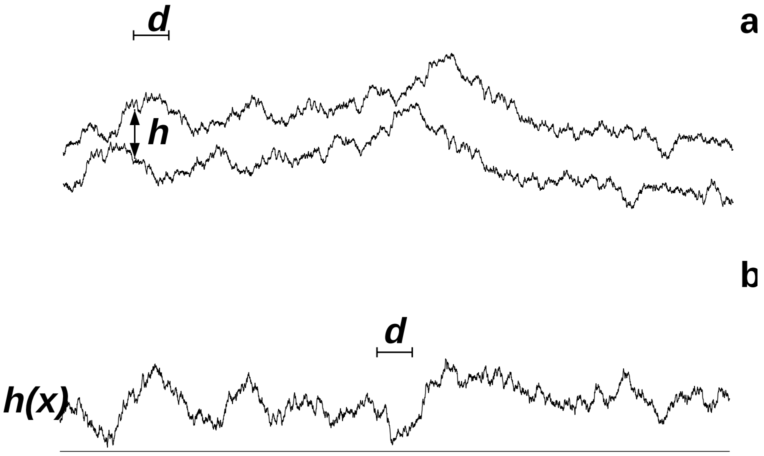

Figure 2.6 Roughness generated by slip on a self-similar surface. (a) The top surface slips a distance d relative to the bottom one. (b) The gap height function $h(x)$ after slip. The characteristic wavelength is comparable to the scale bar of length d.

where K is a constant aspect ratio, ϕ_n is a random phase lag between 0 and 2π, $n_{\max}$ is the maximum value of the index n and λ is the length of the computational region. Since *Power and Tullis* [1991] observed that K is constant for wavelengths from 10 μm to 40 m, the expression in equation (2.25) is valid as long as λ/n is within this range.

We assume that the two rough sides of the fault are initially well-matched before the earthquake due to interseismic processes and the gap height function $h(x)$ is constant along strike. If the fault fluid is a slurry, the thickness of this initially uniform fluid layer is controlled by the fluidization process as well as the fault surface geometry. As one side of the fault slips relative to the other, a mismatch between the fault sides develops and $h(x)$ varies with x (Figure 2.6). Since the initial roughness $f(x)$ is self-similar, the characteristic wavelength of $h(x)$ scales as the displacement between the two fault surfaces. Figure 2.6 shows how initially conformal, rough surfaces result in a high frequency gap height function $h(x)$ when one plane is slipped a distance d relative to the other plane. Larger asperities are isolated with increasing displacements. Visual inspection suggests that d is the dominant asperity length and therefore is approximately equal to L.

A Monte Carlo simulation verifies the scaling relationship between d and L. A series of 300 calculations were performed with an initial gap height function

$$h(x) = f(x + d) - f(x) + h_0 \tag{2.26}$$

where $f(x)$ was randomly generated following equation (2.25). The full pressure distribution and elastic deformation was calculated numerically following the method in Appendix C. The ratio of the P_L computed with equation (2.11) and $L = d$ to the mean fluid pressure from the full calculation measures the accuracy of this scaling relationship. The median of this ratio for convergent trials is 0.75 and 74% of the results are between 0 and 2. The results confirm that the lubrication pressure scales as d. For the purposes of the order of magnitude calculations in this paper $L \approx d$.

About 46% of the trials were non-convergent. Convergence fails in the numerical calculations in some cases due to the inherent instability in the elastohydrodynamic coupling (Appendix C). Other cases fail to converge because the initial geometry diverges too strongly at some point on the fault for lubrication to produce positive pressures with this set of parameters ($K = 10^{-3}$, λ=10 m, d=1 m, η=10 Pa s, h_0=1 mm, $U = 1$ m/s). It is expected that some areas of a real fault surface will fall into this category and therefore be unlubricated. However, most of the trials were entirely lubricated over the full 10 m surface. Even the trials that failed had significant sections of the fault that were favorably oriented for lubrication. The calculations show that for typical geometries and fluid parameters most regions will become lubricated.

Dynamic control of roughness

Figure 2.7 is a calculation of elastohydrodynamic lubrication using a rough surface generated by equation (2.25). The gap height function is produced by slipping this surface a distance d relative to a well-matched surface as in equation (2.26). The increase in deformation with increasing d reproduces the behavior shown in Figure 2.4 for a single asperity with increasing L. The separation between the original and deformed surfaces increases with d. An additional complication in the rough surface

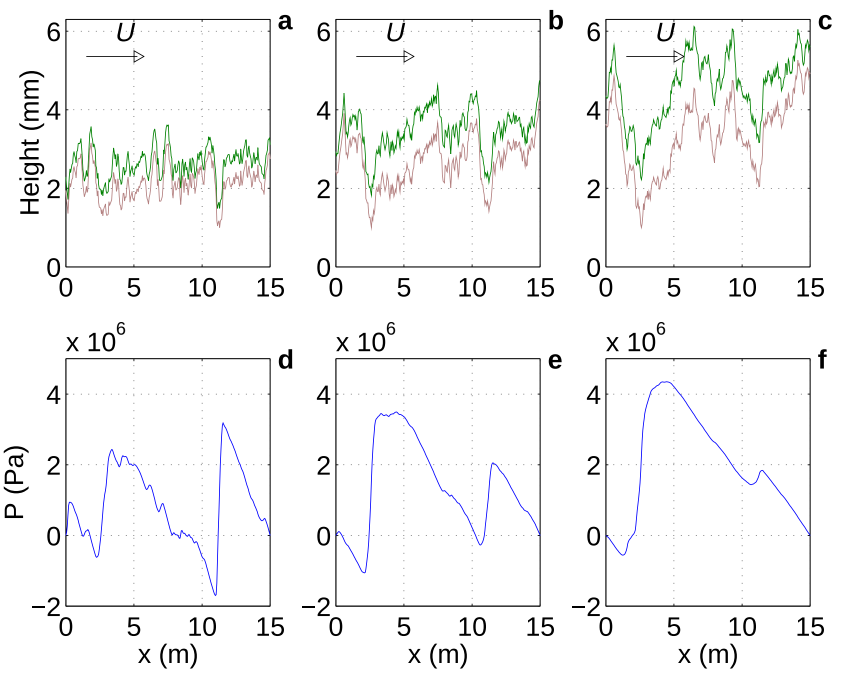

Figure 2.7 Lubrication of a rough surface. (a) The light curve is the undeformed asperity shape and the dark curve is the equilibrium shape. Parameters are as in Table 2.1. $d = 0.5$ m (b) The same as (a) with $d = 1$ m. (c) The same as (a) with $d = 2$ m. (d) The change in pressure in the fluid due to lubrication effects with $d = 0.5$ m.(e) Same as (d) with $d = 1$ m. (f) Same as (d) with $d = 2$ m.

calculations is that the initial asperity height also increases with increasing L. Larger asperities are included in the gap height function as the slip distance increases. Therefore, the flattening is not apparent as in Figure 2.4. If d exceeds the critical value L_c the gap is widened by an amount comparable to the largest asperity heights. The wider gap reduces the number of asperity collisions during sliding. It will be hypothesized that this process has observable effects in nearfield records of high-frequency ground motion.

Physical constraints on parameters

We now use the above formalism to quantify the support provided by lubrication pressure for parameters appropriate to an earthquake.

The value of viscosity depends on the fault fluid. A number of studies have shown that crustal fluids migrate to highly permeable fault zones [*Davis and DeWiest*, 1966] and there is reason to believe that aqueous fluids are abundant at mid-crustal levels due to metamorphic reactions [e.g., *Rumble*, 1994]. If the lubricant during an earthquake is H_2O, the viscosity is on the order of 10^{-3} Pa s [*Sengers and Watson*, 1986]. Alternatively, solid-solid frictional stresses may melt the wall rock during the early stages of an earthquake as first suggested by *Jeffreys* [1942]. Such a melt would be a viable lubricating fluid with a viscosity of at least 10 Pa s [*Spray*, 1993]. In this case thermal effects should be considered in addition to the strictly mechanical lubrication addressed here. We favor a slurry formed from the mixture of fine grained gouge with H_2O as the lubricant for the reasons noted in the introduction. There is evidence of slurries in fault zones and they are probably viscous enough to produce significant lubrication. *Major and Pierson* [1992] showed that at shear strain rates greater than 5 s^{-1} and atmospheric pressure, fine-grained slurries exceed their yield strengths and behave as Newtonian fluids with constant effective viscosities. Since the shear strain rates during rupture are on the order of 1000 s^{-1}, we expect the fluid to be in this Newtonian regime even if some increase of yield strength occurs with confining pressure. The observation of eddy structures in fault zone cataclasites gives further evidence that gouge fluidizes during rupture [*Otsuki*, 1999]. Slurries with mean grain diameters less than 63 μm had viscosities of 0.2 Pa s to 7.1 Pa s in the experiments of *Major and Pierson* [1992] at atmospheric pressure. We expect the viscosity to increase somewhat at depth and therefore estimate the slurry viscosity to be 10 Pa s. *Major and Pierson* [1992] also showed that slurry viscosity is extremely sensitive to solid volume fraction and can change by an order of magnitude if the solid volume fraction changes from 45% to 50%. As a result our estimate of viscosity is necessarily approximate and we acknowledge that it may be inaccurate by a factor of 10 or more.

An average value for slip velocity for moderate earthquakes is 1 m/s [*Lay and Wallace*, 1995]. This standard value is adopted for most of this study. A preliminary model incorporating variable velocity is also explored in this paper under the discussion of radiated energy data as a model application. This more complete treatment includes velocity variations consistent with the variations in friction due to lubrication.

The aspect ratio K is observed to be 10^{-4}–10^{-2} for asperities in the slip parallel direction [*Power and Tullis*, 1991]. In the absence of more precise data, we select $K = 10^{-3}$ as a representative value.

The lubrication length L is approximately equal to the slip distance d as discussed above. The slip distance d scales as $M_0^{1/3}$ where M_0 is the seismic moment [*Kanamori and Anderson*, 1975]. We assume that the slip for a $M_w = 6$ earthquake is 1 m. This approach uses the final slip distance as L and therefore calculates the maximum value of S for a given magnitude earthquake.

The average initial gap height H_0 is assumed to be comparable to observed geological features that are inferred to have been generated during a single slip event. Such features include intravein septa, slickenslide surfaces and occasionally pseudotachylytes. All of these features often indicate sliding localized to regions on the scale of millimeters [*Sibson*, 1999]. Therefore, we take $H_0 = 1$ mm.

The elastic modulus E is a material parameter of the rock and is 5×10^{10} Pa for a typical granite [*Carmichael*, 1982]. Ambient pressure is estimated as the lithostatic value at the hypocenter. Typical values in the seismogenic zone are 1–2×10^{8} Pa. More precise estimates can be made for specific events by using the hypocentral location.

The parameters listed above are summarized in Table 2.1 and result in values of the Sommerfeld number on the order of 0.1 for moderate earthquakes. These values indicate that the rupture process overlaps with the dynamic regime in which variations in lubrication pressure are significant. Values for a few representative events are listed in Table 2.2.

Combining the above constraints and using the standard relationship between seismic moment and magnitude,

$$\log M_0 = 1.5M_w + 9.1, \tag{2.27}$$

Table 2.1. Typical Parameter Ranges

Parameter	Values
η	10 Pa s
U	1 m/s
K	10^{-3}
L	$(M_0/M_0^0)^{1/3}$ m[a]
H_0	10^{-3} m
E	5×10^{10} Pa
P	1–2 × 10^8 Pa

[a] M_0^0 is the moment of a $M_w = 6$ earthquake (1.3 $\times 10^{18}$ N m) and is used to normalize the slip distance to be 1 m for such an event.

Table 2.2. Representative Events and Source Parameters

Earthquake	M_w	U m/s	d m	P 10^8 Pa	S
Landers[1] 6/28/92	7.2	1	7	1.9	0.1
Northridge[2] 1/17/94	6.7	1	3.2	4.7	0.1
Parkfield[3] 12/20/94	4.7	1.1	0.77	2.4	0.1

Sommerfeld numbers for representative earthquake. Values for P are based on hypocentral depths. In order to provide consistent comparisons, Sommerfeld number is computed for the maximum observed velocity and slip. Parameters omitted from this table are taken to be the typical values in Table 2.1. References: [1] *Wald and Heaton* [1994]; [2] *Wald et al.* [1996]; [3] *Fletcher and Spudich* [1998].

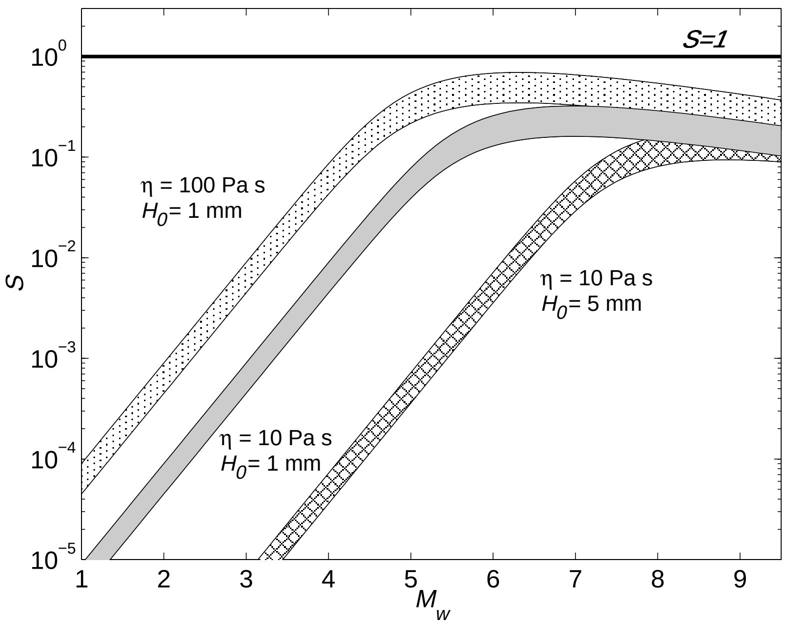

Figure 2.8 Dependence of Sommerfeld number on magnitude M_w. The solid filled area is calculated withe the typical parameters in Table 2.1. The upper and lower bounds are determined by the range in lithostatic pressure P. The other two regions show the sensitivity of the results to varying parameters. Both are also calculated with the parameters of Table 2.1 except η=100 Pa s in the stippled area and H_0=5 mm in the cross-hatched area.

where M_0 has units of N m, we graph a relationship between Sommerfeld number and magnitude (Figure 2.8). Lubrication pressure can reach 30% of the lithostatic pressure for large earthquakes with typical fault parameters, but lubrication is negligible for small earthquakes.

Parameters in Table 2.1 are approximate and Figure 2.8 shows the sensitivity of the results to errors in the estimates. The least constrained parameters, η and H_0, are varied. The results are very sensitive to the parameters chosen for small earthquakes and more robust for large earthquakes. Despite the variation, small earthquakes remain essentially unlubricated in all cases. The significant lubrication pressures in moderate to large events are relatively insensitive to variations in parameters because of the limiting effect of the elastic deformation. This result is consistent with the scaling argument in equation (2.17) where the Sommerfeld number is shown to be proportional to the 1/4 root of most parameters.

Model applications

The previous sections presented a conceptual overview of hydrodynamic lubrication, a mathematical model and the necessary quantification of parameters. We now use lubrication theory to explain the following three observations: low heat flow on the San Andreas, along-strike variations in strong motion spectra and a difference in scaled radiated energy between large and small earthquakes. At the end of this section we present an optimized parameter set that is consistent with all three applications.

Heat flow

Heat flow measurements across the San Andreas fault require that the frictional stress on the fault is less than 10^7 Pa at 7 km depth [*Lachenbruch and Sass*, 1980]. This strong constraint on the fault energetics is puzzling in light of laboratory data on rock friction. According to laboratory experiments, the frictional stress τ_f between rock surfaces is determined by Byerlee's Law

$$\tau_f = \mu_s P_e = \mu_s (P - P_p^0 - P_L) \tag{2.28}$$

where the coefficient of friction μ_s is between 0.6 and 0.85 and the effective pressure is as discussed for equation (2.18). At an average seismogenic depth of 7 km with hydrostatic fluid pressure ($P = 1.9 \times 10^8$ Pa, $P_p^0 = 7 \times 10^7$ Pa), Byerlee's law requires the frictional stress to be at least 7×10^7 Pa (Table 2.3). A dynamic increase in fluid pressure of 10^8 Pa, i.e., $S = 0.5$ is required in order to solve the friction paradox. The parameters considered in Table 2.1 allow a maximum lubrication pressure $P_L = 3 \times 10^7$ Pa. Dynamic lubrication in this situation reduces the friction by 30% relative to the hydrostatic value, but does not completely solve the friction paradox. Sufficient lubrication pressure can be achieved by lubrication using a higher viscosity

Table 2.3. Frictional Stress

Case	P_p^0 (10^7 Pa)	P_L (10^7 Pa)	τ_f (10^7 Pa)
a	0	0	10
b	7	0	7
c	7	3	5
d	7	10	0.9

Minimum frictional stress on the fault τ_f computed from Byerlee's law (2.28) for different assumptions. The cases are: (a) Dry rock. (b) Hydrostatic pore pressure only. (c) Lubrication with typical parameters. (d) Lubrication with η=350 Pa s.

slurry or melt with $\eta = 350$ Pa s. Alternatively, other mechanisms such as thermal pressurization in the confined fault zone might account for a portion of the dynamic fluid pressure increase [*Sibson*, 1973; *Lachenbruch*, 1980; *Mase and Smith*, 1987].

The calculations show that hydrodynamic lubrication is only significant for large earthquakes. A dynamic lubrication model for reducing frictional stress requires that the majority of slip on a fault occurs during the larger magnitude events. This is consistent with standard scaling relationships for seismic energy with magnitude [*Kanamori and Anderson*, 1975].

Strong motion spectra

A unique set of strong motion records was recorded within a few kilometers of the surface rupture during the September 20, 1999, $M_w = 7.6$ Chi-Chi, Taiwan, earthquake (Figure 2.9). The earthquake is notable for its high particle velocities as well as the distinct lack of high frequency energy in the areas with large velocities [*Ma et al.*, 2000]. The northern records are dominated by a smooth, large amplitude pulse in velocity, whereas the southern ones have higher frequency energy with lower velocities. This difference in frequency content along-strike was also manifested as greater damage in the south, despite the larger displacements in the north. The damage patterns and mapped fault offsets confirm that the variation in the records is a re-

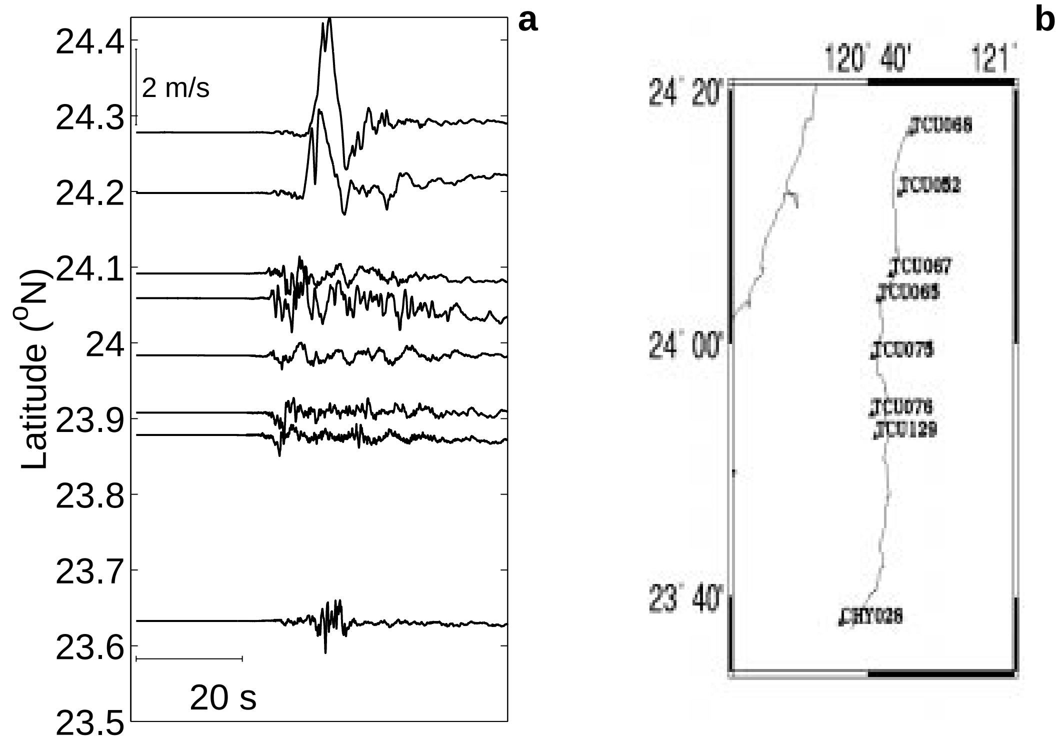

Figure 2.9 Velocity records from Chi-Chi, Taiwan $M_w = 7.6$ earthquake. (a) North component of velocity at stations within a few kilometers of the surface rupture. Velocity is integrated from acceleration records provided courtesy of Taiwan Meteorological Institute with a correction for instrumental drift late in the record. The y-axis shows latitude of the stations. (b) Station map and surface rupture. Fault plane dips to the East.

sult of along-strike trends rather than local site responses or amplified motion on the hanging wall.

We hypothesize that the high frequency energy is generated by contacting asperities (Figure 2.10). Collisions between asperities are accommodated by a combination of normal displacement and failure. Both processes radiate high frequency energy. Asperities that radiate energy by these processes must be much smaller than the slip distance. The rate of asperity contact controls the minimum frequency, f_c, of the displacement waves generated by asperity collisions. Therefore,

$$f_c \gg U/d. \tag{2.29}$$

Using the values in Table 2.3 we compute that for station TCU129, $f_c \gg 0.8$ Hz and for TCU068, $f_c \gg 0.4$ Hz. Based on these values we define "high frequency energy" as waves with frequencies greater than 1 Hz.

The lubrication model can explain the observed data if the lubrication lengthscale L in the northern section of the fault was larger than the critical value L_c and therefore elastic deformation was significant. For the self-similar rough fault L is identical to the slip distance d as shown in Figure 2.6. It is hypothesized that at the northern stations $d \gg L_c$; therefore, the elastic deformation widened the fault and reduced asperity contact. High frequency radiation was reduced. Elsewhere on the fault d is hypothesized to have been less than L_c throughout the earthquake.

We verify the consistency of the hypothesis with the data by comparing the observed maximum slips (L) with the theoretical values of L_c from equation (2.16). The results calculated with the standard lubrication parameters in Table 2.1 combined with the velocities and slip distances in Table 2.3 are shown in Figure 2.11. The critical slip is between 1 and 2 m at all of the stations. L is shown to exceed L_c only at the two northernmost stations and therefore the results are consistent with the lubrication hyopthesis. The values of L_c presented here depend on the estimates of the parameters in Table 2.1. However, equation (2.16) shows that the results are relatively insensitive to errors in all the parameters except H_0. The critical slip L_c is

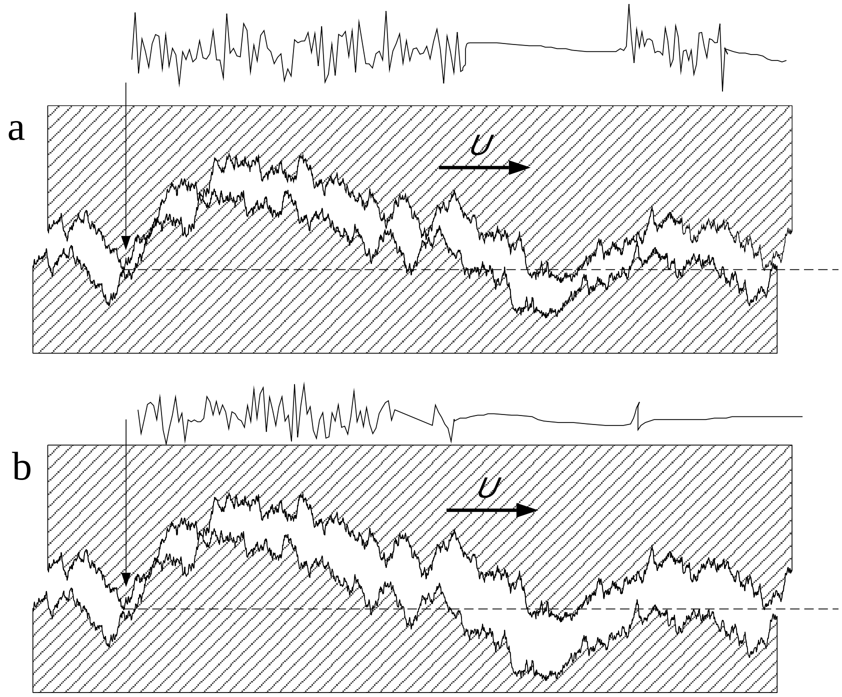

Figure 2.10 Cartoon illustrating asperities breaking during rupture. (a) Undeformed fault surface. As the point at the tip of the arrow slips to the right, it collides with the asperities on the lower surface above the dotted line. A schematic high-pass velocity trace for a station located at the arrow is shown above the figure. (b) Deformed fault surface. There are fewer asperity contacts than in the undeformed case as shown by the shading and the velocity trace.

Table 2.3. Taiwan Station Parameters

Station	d (m)	U (m/s)
TCU068	9.6	3.9
TCU052	8.7	2.7
TCU067	1.2	1.1
TCU065	1.6	1.5
TCU075	1.3	1.2
TCU076	1.3	0.87
TCU129	1.0	0.78
CHY028	1.3	0.94

Observed maximum horizontal slip and particle velocity. Stations are arranged north to south as in Figure 2.9. The first two stations show anomalous “pulse” behavior.

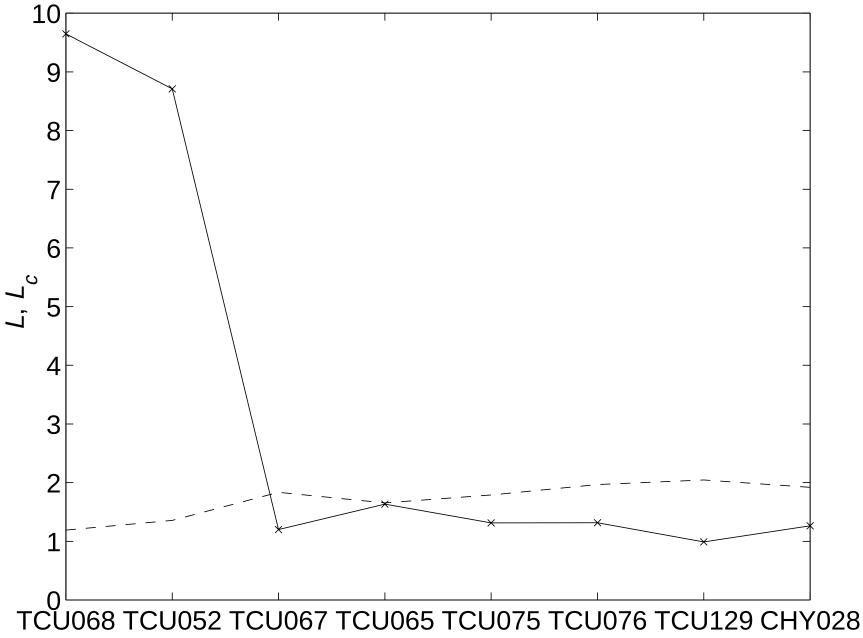

Figure 2.11 Comparison of observed maximum horizontal slip ($L = d$) at each station as tabulated in Table 2.3 (solid line) with the values of L_c computed using equation (2.16) (dashed line).

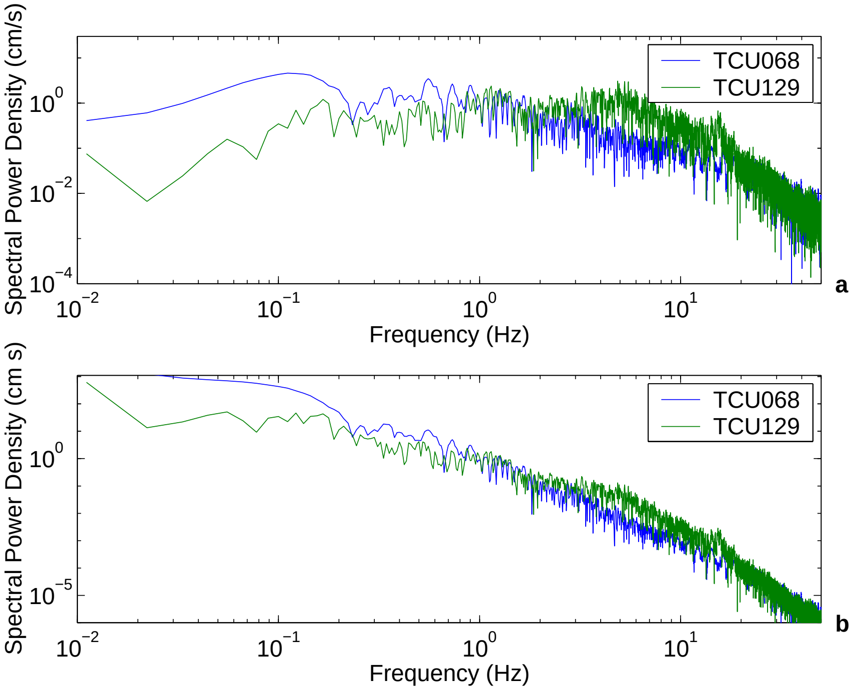

Figure 2.12 (a) Acceleration spectra for the Chi-Chi earthquake of the N component from a northern (TCU068) and southern (TCU129) near-fault station. (b) Displacement spectra derived from the acceleration records.

likely to be of the order of a few meters. Spectral analysis of the northern stations shows a fall-off in frequency above ∼1 Hz as would be expected from the lubrication model (Figure 2.12).

We predict that any large earthquake with slip distances greater than a few meters will have a zone of the fault that is well-lubricated with depleted high-frequency energy. This behavior has direct implications for efforts to mitigate the effects of strong ground motion. Structures designed to withstand large earthquakes must accommodate large amplitude long-period motion, but do not need to accommodate as much high frequency energy as would be predicted from scaling the spectrum of small events. Of course, such theoretical predictions are preliminary and should be thoroughly validated before being incorporated into any design plans.

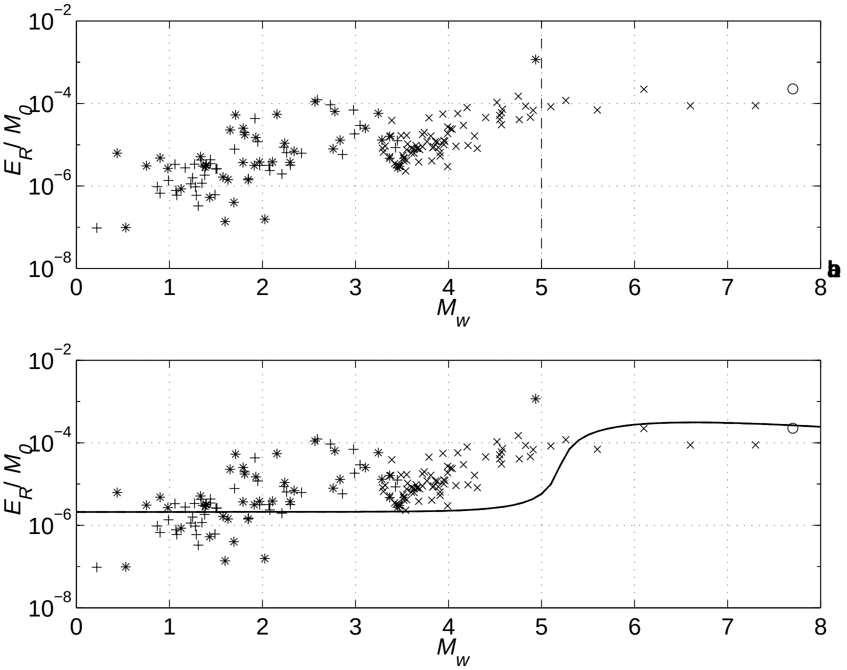

Figure 2.13 Observations of the ratio of locally measured radiated energy E_R to M_0. Data is from the following sources: stars, *Prejean and Ellsworth* [submitted]; pluses, *Abercrombie* [1995]; x's *Kanamori and Heaton* [2000]; circle, *Ma et al.* [in prep.]. (a) The data is separated into two regimes. To the right of the dashed line E_R/M_0 is approximately constant and to the left it gradually decreases. Scatter also increases for the small events as discussed in the text. (b) The data and a solid line indicating the modeled values of E_R/M_0 using (2.12) and (2.30)–(2.33). We only attempt to fit the change in E_R/M_0 by a factor of 100 for M_w=4 to M_w=6. Parameters are as in Table 2.1 and as follows: $G = 2.0 \times 10^{10}$ Pa, $\mu_s = 0.6$, $P = 10^8$ Pa, $\sigma_0 = \mu_s P + 2 \times 10^5$ Pa, $V_r/\beta = 0.65$, $\beta = 3300$ m/s.

Radiated energy

It has been observed in California that the ratio of radiated energy E_R to seismic moment M_0 is a function of magnitude (Figure 2.13) [*Abercrombie*, 1995; *Prejean and Ellsworth*, submitted; *Kanamori and Heaton*, 2000]. Large events ($M_w > 6$) have values of $E_R/M_0 \approx 10^{-4}$ while the small ones ($M_w < 4$) have values as low as 10^{-6}. This type of observation is extremely difficult and the error increases for small events because of the model-dependent corrections for attenuation. Such corrections are necessary for accurate measurement of radiated energy and likely obscure the systematics for small earthquakes. We therefore only attempt to interpret the overall trend. There is a constant, large value of E_R/M_0 for large events and a transition occurs at $M_w \approx 5$ to smaller values of the ratio for small earthquakes. We attempt to fit only the observed change of E_R/M_0 by a factor of 100 from M_w=4 to M_w=6.

The observations of scaled radiated energy have particularly important implications for slip velocity and therefore frictional behavior. Both *Sato and Hirasawa* [1973] and *Mott* [1948] derived a relationship between rupture velocity V_r, driving stress $\Delta\sigma_d$ and the observable ratio E_R/M_0, by calculating the kinetic energy during rupture.

$$\frac{E_R}{M_0} = \frac{\Delta\sigma_d}{2G}\left(\frac{V_r}{\beta}\right)^2 \tag{2.30}$$

where G is the shear modulus and β is the shear velocity in the country rock outside the fault zone. The average driving stress during an earthquake is related to the initial stress σ_0 and the frictional stress σ_f by

$$\Delta\sigma_d = \sigma_0 - \sigma_f \tag{2.31}$$

and the frictional stress is reduced by the lubrication pressure. Therefore, the driving stress is

$$\Delta\sigma_d = \sigma_0 - \mu_s P + \mu_s P_L. \tag{2.32}$$

The lubrication pressure P_L depends on the slip velocity U as shown in (2.11). The

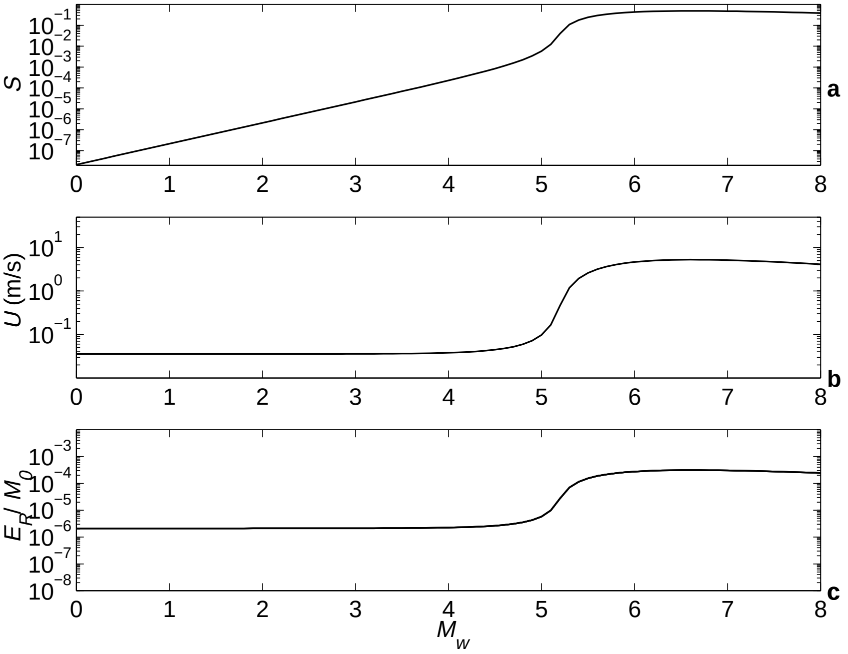

Figure 2.14 Dynamic friction model results. Frictionally controlled (a) Sommerfeld number, (b) slip velocity and (c) radiated energy as a function of magnitude. Parameters are as in Figure 2.13.

relative slip velocity is related to the driving stress in a propagating shear crack by

$$U = \Delta\sigma_d \beta / G \frac{2V_r/\beta}{E\left[\sqrt{1-(V_r/\beta)^2}\right]} \tag{2.33}$$

where $E[\,]$ is the complete elliptic integral of the second kind [*Husseini*, 1977]. The derivation of (2.33) assumes that the driving stress is equal to the static stress drop. Here we assume that the equation is equally applicable for the instantaneous driving stress σ_d.

Combining (2.32), (2.33) and the definition of P_L in (2.11), we solve for U. The model nonlinearly couples the driving stress, slip velocity and lubrication pressure. In Figure 2.13 solutions of S and U as a function of M_w are computed and then E_R/M_0 is derived using (2.30). The behavior of this model is best understood in terms of the critical slip distance L_c defined by equation (2.16). For the inelastic regime where $L < L_c$, as the slip distance increases, the lubrication pressure rises. For small magnitudes ($M_w < 4$), the lubrication pressure is much less than lithostatic pressure ($S \ll 1$) and there is little effect on the total frictional stress. Therefore, velocity and E_R/M_0 are nearly constant in this magnitude range. At $M_w \approx 4$ the lubrication pressure becomes a significant fraction of the lithostatic pressure and the frictional stress is reduced. Since the frictional resistance is less than in the unlubricated case, the driving stress is increased. The driving stress is coupled to U and E_R/M_0 in equations (2.30) and (2.33). Therefore, U and E_R/M_0 also increase relative to their values for small events ($M_w < 4$).

For large magnitude earthquakes ($M_w > 6$) where $L > L_c$ and elastic deformation is significant, the lubrication pressure decreases with increasing magnitude. The lubrication pressure P_L gradually decreases as $L^{-1/4}$ for large events in accordance with equation (2.17). The decrease in lubrication pressure results in a small increase in frictional resistance and therefore a decrease in the available driving stress. As a result, the slopes of the velocity and radiated energy curves in Figure 2.14 are slightly negative. At large magnitudes the model predicts a gradual decrease in radiated energy with increasing magnitude. This decrease is not observed, but the trend is much smaller than the scatter in the data. One could infer that this inconsistency is due to the sparse data set, or that it is the result of the specific assumptions in the model. For instance, it is possible that V_r/β is not constant for all magnitudes. As the proportion of fracture energy in the total energy balance decreases for large earthquakes, the relative rupture velocity increases. Including a linear increase in V_r/β would improve the fit of the model, but invokes physics that is beyond the scope of the theory presented here.

We noted above that the increase in scatter for small earthquakes may be due to the methods used to measure radiated energy for small events. We speculate that there may also be real variability in E_R/M_0 for small events due to variations in parameters in the unlubricated regime. In the lubricated regime the system is insensitive to differences in the fault properties as shown in Figure 2.8 and equation (2.17). The scatter in E_R/M_0 data for large earthquakes is therefore small.

The earlier calculation of Sommerfeld number versus magnitude in Figure 2.8 appears to be superseded by Figure 2.14. However, the latter figure contains more model dependent assumptions embedded in (2.30)–(2.33). We have therefore chosen to present both forms of the relationship between M_w and Sommerfeld number S.

Optimized parameters

We present a heuristically optimized set of parameters that fits all three model applications in Table 2.4. The approach taken in this section is to find parameters that match all of the datasets reasonably well instead of assigning the typical values of Table 2.1 *a priori*. No formal optimization algorithm was used since the fit is nonunique. The values in Table 2.4 are merely one representative set that can match all three datasets. The major difference between the optimized and typical parameter sets is that the viscosity is larger in the optimized set by a factor of 5. Since small changes of the particle concentration can change the slurry viscosity by an order of magnitude, we consider this adaptation reasonable.

All three model applications demonstrate the most important prediction of elastohydrodynamic lubrication theory. Large earthquakes behave qualitatively differently than small ones. Events with slip distances greater than ~0.5 m have low dynamic friction, fewer asperity-asperity contacts and a larger proportion of the energy radiated.

Discussion

Now that the magnitude of the hydrodynamic lubrication effect is calculated, it is appropriate to review the assumptions about the fluid dynamics that were incorporated into the model. It is assumed that the fluid forms a continuous layer of constant initial height in the gap between moving, subparallel surfaces. This fluid may be a slurry which forms during the initiation of the earthquake. The duration of the earthquake is assumed to be much longer than the time necessary for fluidization. In the lubricating layer a continuum approximation is assumed to be appropriate. The

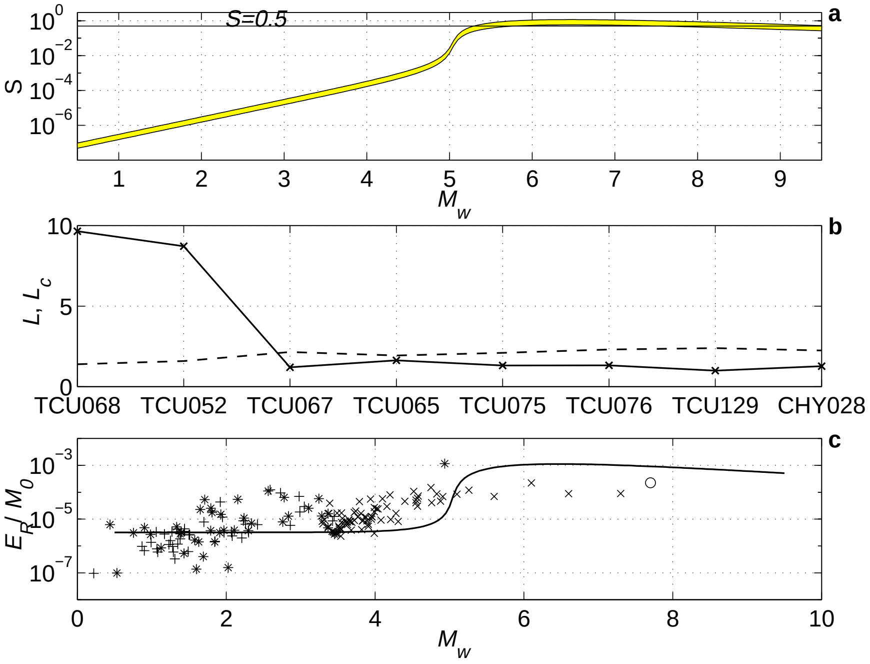

Figure 2.15 All three model applications with an optimally selected parameters set from Table 2.4. (a) Sommerfeld number as a function of M_w used to match heat flow data. S is calculated using the model of equations (2.32)–(2.33). (b) Observed lubrication length at each station compared to the theoretical critical value computed using the optimized parameters and the observed maximum velocities in Equation (2.16). (c) Predicted E_R/M_0 as a function of M_w and observed values. E_R/M_0 changes by a factor of 300 from M_w=4 to M_w=6.

Table 2.4. Optimized Parameters

Parameter	Values
η	50 Pa s
U	see (2.33)
K	2×10^{-3}
L	$(M_0/M_0^0)^{1/3}$ m
H_0	2×10^{-3} m
E	5×10^{10} Pa
P	1×10^{8} Pa
G	2.0×10^{10} Pa
μ_s	0.6
σ_0	6.01×10^{7} Pa
V_r/β	0.8
β	3.3 km/s

sediment particles in the slurry are much smaller than the gap height. In the farfield where there is no slip, the pressure remains at the pre-earthquake level. The transition from the slipping to non-slipping zone is not explicitly modeled and is assumed to have only local effects.

Additional complications could be added to the model and are likely to perturb the results presented here. The purpose of this study is to present a new concept in fault mechanics along with the minimum necessary quantification. Further refinements are clearly possible, but here we confine ourselves to only a brief evaluation of a few of the possible complicating factors.

In the lubricated fault, fluid flows into a narrow gap and the pore pressure dynamically increases. The permeability of the surrounding rock is sufficiently low that the fluid is effectively confined by the walls. This model is clearly an idealized end-member. In the natural system some leakage occurs during faulting to reduce the lubrication pressure. Such reductions can be included numerically as they are in the porous bearing of *Kaneko et al.* [1999] or the thermal pressurization calculations of *Lachenbruch* [1980]. Note that the fluid pressures calculated for the typical parameters of Table 2.1 are at most 40% lithostatic and therefore insufficient to produce hydrofractures as a means of releasing pressure. The optimized parameter set produces higher pressures and may result in hydrofracture.

Common lubricants in engineering applications have pressure dependent viscosities where the viscosity increases with pressure. This has the effect of increasing the lubrication pressure as there is a positive feedback in the process. Similar behavior might be expected for fault slurry, but silicate melt viscosity decreases with pressure [*Kushiro*, 1980]. Therefore, any identification of the fault fluid as melt must carefully consider whether the pressure dependent viscosity would eliminate the lubrication effect.

The two-dimensional fault model neglects flow perpendicular to the slip direction. Such "side-leakage" for a lubricated gap with equal length and width can reduce the fluid gap height by a factor of 2–3 for a constant load [*Hamrock*, 1994, p. 483]. It is not clear what the effect would be in the case of a fault where the fluid layer height is constant and the load varies. In addition, the contacting asperities set much more complex boundary conditions than are addressed here. A fully coupled solid-fluid model is required for a rigorous and complete study. The theoretical and numerical underpinnings of such models are at the forefront of current research in tribology [*Spikes*, 1997; *Jiang et al.*, 1999].

The normal velocity of the wall, V, is neglected in the form of the Reynolds equation used here. This approximation is valid if $V \ll UH/L$ [*Leal*, 1992]. The quasistatic model used for the elastic deformation also assumes that V is small therefore the model is self-consistent. However, a fully dynamic calculation may show that the wall velocity is important in modifying the lubrication pressure.

In bearings, the lubricating fluid is seldom able to support negative dynamic pressures [*Dowson and Higginson*, 1977]. A combination of cavitation and exsolution of dissolved gases maintains the pressure at the initial reservoir level in strongly diverging channels. Cavitation is not expected to occur on a fault with a large confining pressure, but exsolution may be possible. If exsolution occurs and prevents negative dynamic pressures, then the net effect of hydrodynamic lubrication is much greater than calculated here.

Another possible complication is the compressibility of the fluid. Since the lubrication pressure is at most of the order of the initial hydrostatic pressure, the effects of compressibility on the pressure distribution are expected to be small [*Szeri*, 1998]. *Dowson and Higginson* [1977] showed for rectangular elastohydrodynamic conjunctions that fluid compressibility has no significant effect on the elastic deformation although the pressure distribution is slightly different than in the incompressible case. Preliminary numerical experiments suggest that these results are also applicable to the rough fault studied here, but more work would be necessary to fully exclude the importance of compressibility.

As the fluid is squeezed through the narrow gap, dissipative heating occurs that may increase the temperature of the system. This thermal effect could decrease lubrication by decreasing the viscosity. At the same time, the thermal expansion of the fluid would increase the fluid pressure. Additional complications could arise if the fluid is Non-Newtonian and the viscosity depended upon the shear rate [*Jacobson*, 1991]. Inelastic effects such as wear of the wall rock may also be important.

Conclusions

We have shown in this paper that the mechanical effect of a viscous fluid lubricating a fault zone has implications for the rupture dynamics. Lubrication reduces the frictional stress during an earthquake by as much as 50% for typical parameters and can completely eliminate it if the fault fluid has a viscosity of 50 Pa s. Dynamic widening of the fault reduces the radiation of high-frequency (>1 Hz) energy for earthquakes with large slips. The two order of magnitude difference in the ratio of radiated energy to moment between small and large earthquakes is interpreted to reflect the lubrication-controlled frictional properties. All of these applications utilize the variation in lubrication behavior between small and large slip events. This change in behavior at the critical slip distance L_c is one of the most intriguing consequences of this theory and warrants further study.

Hydrodynamic lubrication does not provide a complete answer to the "friction paradox." However, the mechanism outlined here indicates that the static values of pore pressure are not appropriate for studies of earthquake rupture. The lubrication effect is present in all fluid-filled faults and is independent of any assumptions about the thermal effects of fault friction. The pore pressure must be viewed as a dynamic quantity and the pre-rupture values should not be used for modeling coseismic friction.

Appendix A: The lubrication approximation

The standard simplification of the Navier-Stokes equation for a thin film follows. For more detailed discussions see *Szeri* [1998] or *Leal* [1992].

The fluid motion is completely described by the full Navier-Stokes equation,

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla p + \eta \nabla^2 \mathbf{u}, \tag{2.34}$$

and the continuity equation

$$\rho \nabla \cdot \mathbf{u} = \frac{D\rho}{Dt} \tag{2.35}$$

where ρ is the density, $\mathbf{u}$ is the velocity vector, p is the pressure and η is the viscosity. In a two-dimensional, incompressible flow,

$$\rho\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial z}\right) = -\frac{\partial p}{\partial x} + \eta\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}\right) \tag{2.36}$$

$$\rho\left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial z}\right) = -\frac{\partial p}{\partial z} + \eta\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2}\right) \tag{2.37}$$

and

$$\frac{\partial u}{\partial z} + \frac{\partial v}{\partial z} = 0 \tag{2.38}$$

where u is the velocity in the x direction and v is the velocity in the z direction. We adopt a coordinate system in which x is defined as parallel to the slip vector and z is the normal to the fault plane. For a narrow slit geometry, such as would be commonly found in a fault plane, some simplifications can be made. The length L over which the velocity varies significantly in the x direction is much greater than the distance H over which velocities vary in the z direction (Figure 2.2). This separation of characteristic lengths for the two dimensions results in a natural separation of the terms according to the power of H/L. This result can be derived by replacing the variables by their non-dimensional equivalents as follows:

$$\bar{x} \equiv x/L \tag{2.39}$$

$$\bar{z} \equiv z/H \tag{2.40}$$

$$\bar{u} \equiv u/U \tag{2.41}$$

$$\bar{v} \equiv v/(UH/L) \tag{2.42}$$

$$\bar{p} \equiv p/(\eta UL/H^2) \tag{2.43}$$

$$\bar{t} \equiv t/(L/U) \tag{2.44}$$

where U is the relative velocity of the two sides of the fault. The above scaling uses the continuity equation to establish the characteristic scale for v and the Navier-Stokes equation to find the characteristic value of p. Equations 2.36 and 2.37 in terms of the above non-dimensional variables are

$$\frac{\rho UL}{\eta}\left(\frac{H}{L}\right)^2\left(\frac{\partial \bar{u}}{\partial \bar{t}} + \bar{u}\frac{\partial \bar{u}}{\partial \bar{x}} + \bar{v}\frac{\partial \bar{u}}{\partial \bar{z}}\right) = -\frac{\partial \bar{p}}{\partial \bar{x}} + \left(\frac{H}{L}\right)^2\frac{\partial^2 \bar{u}}{\partial \bar{x}^2} + \frac{\partial^2 \bar{u}}{\partial \bar{z}^2} \tag{2.45}$$

$$\frac{\rho UL}{\eta}\left(\frac{H}{L}\right)^4\left(\frac{\partial \bar{v}}{\partial \bar{t}} + \bar{u}\frac{\partial \bar{v}}{\partial \bar{x}} + \bar{v}\frac{\partial \bar{v}}{\partial \bar{z}}\right) = -\frac{\partial \bar{p}}{\partial \bar{z}} + \left(\frac{H}{L}\right)^4\frac{\partial^2 \bar{v}}{\partial \bar{x}^2} + \left(\frac{H}{L}\right)^2\frac{\partial^2 \bar{v}}{\partial \bar{z}^2}. \tag{2.46}$$

The thin gap geometry implies that $H/L \ll 1$. The Reynolds number $Re \equiv \rho UL/\eta$ appears on the left-hand side of these equations. If $Re \ll (L/H)^2$, the inertial forces are negligible. To first order in H/L, the Navier-Stokes equations are:

$$\frac{\partial p}{\partial x} = \eta\frac{\partial^2 u}{\partial z^2} \tag{2.47}$$

$$\frac{\partial p}{\partial z} = 0. \tag{2.48}$$

Equations (2.47) and (2.48) are the equations of motion for a lubricating layer. Pressure gradients across the gap are negligible and the dynamics are dominated by the

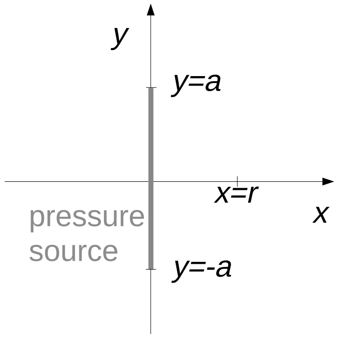

Figure 2.16 Geometry for calculating deformation from an finite line source of pressure. In the geological case, the $x - y$ plane is the fault plane and slip is in the x direction.

balance of the viscous stresses and the pressure gradient parallel to the gap.

Appendix B: Elastic deformation solution

We begin with the solution for the deformation from a force of magnitude F applied at the origin on the free surface of an elastic halfspace [*Timoshenko and Goodier*, 1970],

$$w(r) = \frac{F(1-\nu^2)}{\pi E r}, \tag{2.49}$$

where $w(r)$ is the vertical deformation on the free surface at a distance r from the origin and E is Young's modulus. This is equivalent to a pressure source $p(x, y)$ applied at a point,

$$p(x, y)\delta_d(x)\delta_d(y) = F \tag{2.50}$$

where $\delta_d()$ denotes a Dirac delta function.

A finite line on the y-axis with ends at $y = \pm a$ generates a deformation that is the sum of the point source terms (Figure 2.16). If the pressure is constant in the y direction, i.e., $\partial p/\partial y = 0$, the displacement at a point on the x-axis that is a distance

r from the origin is

$$w(r) = 2\int_{y=0}^{a} \frac{(1-\nu^2)p(x)}{\pi E r} dy \,. \tag{2.51}$$

If the finite line source is at $x = \xi$, the displacement $w(x)$ in a Cartesian coordinate system is

$$w(x) = 2\delta_d(\xi) \int_{y=0}^{a} \frac{p(\xi)(1-\nu^2)}{\pi E\sqrt{(x-\xi)^2 + y^2}} dy. \tag{2.52}$$

After performing the integration, the displacement is explicitly

$$w(x) = \frac{2p(\xi)(1-\nu^2)}{\pi E}\delta_d(\xi) \ln \left| \frac{a + \sqrt{(x-\xi)^2 + a^2}}{x-\xi} \right| . \tag{2.53}$$

The deformation from a pressurized fault of length L is the integration of (2.53) over every point $x = \xi$. The change in total gap height $\delta(x)$ combines contributions from both walls,

$$\delta(x) = \frac{4(1-\nu^2)}{\pi E} \int_0^L p(\xi) \ln \left| \frac{a + \sqrt{(x-\xi)^2 + a^2}}{x-\xi} \right| d\xi. \tag{2.54}$$

In the model described in the text, we assume that the lubricated zone is equidimensional and therefore $2a = L$.

Appendix C: Numerical method

Incorporating elastic deformation into lubrication in a thin gap requires the simultaneous solution of the Reynolds equation (2.6) and the elastic deformation of the boundaries (2.8). The gap height $h(x)$ is related to the initial gap height $s(x)$ and the elastic deformation $\delta(x)$ by

$$h(x) = s(x) + \delta(x). \tag{2.55}$$

These equations are to be solved with the boundary conditions,

$$p(x = 0) = 0$$

$$p(x = L) = 0. \tag{2.56}$$

Reynolds equation

In the absence of elastic deformation, the Reynolds equation (2.6) is solved by considering its integrated form.

$$p(x) = -6\eta U \int_0^x \frac{h - h^*}{h^3} dx + p_p^0 \tag{2.57}$$

We use the boundary conditions (2.56) to solve for the integration constant h^* in (2.57). As discussed before, the pressures are referenced to the level where $p_p^0 = 0$. The boundary condition at $x = L$ is $p = 0$ which implies that

$$h^* = \frac{\int_0^L \frac{1}{h^2} dx}{\int_0^L \frac{1}{h^3} dx}. \tag{2.58}$$

In order to numerically solve for p we use a discretization of (2.57) utilizing the trapezoidal rule [*Abramowitz and Stegun*, 1965]

$$\tilde{p}_i = -6\eta U \sum_{j=1}^{i-1} \left(\frac{h_j - h^*}{h_j^3} + \frac{h_{j+1} - h^*}{h_{j+1}^3} \right) \frac{\Delta x}{2} \tag{2.59}$$

where $\Delta x \equiv x_{i+1} - x_i$ and $\tilde{p}_i$ is the solution to the Reynolds equation for the pressure at $x = i\Delta x$. In the absence of elastic deformation $h(x) = s(x)$ where $s(x)$ is the initial gap height function.

Elastic deformation

The elastic deformation equation (2.8) is discretized as

$$\delta_i = \frac{4(1-\nu^2)\Delta x}{\pi E} \sum_{j=1}^{N} p_j \ln \left| \frac{L/2 + \sqrt{(L/2)^2 + (x_{i+1/2} - x_j)(x_{i-1/2} - x_j)}}{\sqrt{(x_{i+1/2} - x_j)(x_{i-1/2} - x_j)}} \right| \tag{2.60}$$

where N is the number of grid points. The terms $x_{i+1/2}$ and $x_{i-1/2}$ are introduced following *Okamura* [1982] to handle the singularity where $x_i = x_j$.

The elastic deformation is combined with the Reynolds equation using an iteration scheme with damping. Instead of using the full computed $\tilde{p}(x)$ from (2.59), a portion of the calculated pressure is added at each step. The pressure distribution for step $k+1$ is calculated based on the pressure at step k by

$$p_i^{k+1} = p_i^k + \omega(\tilde{p}_i - p_i^k) \tag{2.61}$$

where ω is a damping factor between 0 and 1. For gaps with an initial average height of 0.1 mm, ω must be 0.01 or less during the first two steps for convergence to be achieved in the majority of cases. After the initial few steps, the problem becomes more stable and the degree of damping can be reduced adaptively. For randomly generated rough surfaces, ω=0.01 is occasionally insufficient and the calculation does not converge. The hydrodynamic code and iteration scheme were verified with analytic and approximate solutions for a planar slider like that shown in Figure 2.3.

The full scheme for computing iteration step k is as follows:

1) An initial gap height $s(x)$ is chosen and calculated as a function of x.

2) The deformed gap height $h(x) = s(x) + \delta(x)$ is calculated. During the first iteration $\delta(x)$ is assumed to be 0 everywhere, so $h(x) = s(x)$.

3) The integrated, discretized Reynolds equation in (2.59) is solved for $\tilde{p}(x)$ using the geometry given by $h(x)$.

4) The new pressure $p^{k+1}(x)$ is calculated using (2.61). Initially $p(x) = \mathbf{0}$; therefore, $p^{k+1}(x) = \omega\tilde{p}(x)$ in the first iteration.

5) The deformation $\delta(x)$ is calculated with (2.60) using the new pressure distribution $p^{k+1}(x)$.

6) The convergence parameter $\text{Dif} \equiv (\tilde{p}_l - p_l^k)/p_l^{k+1}$ is evaluated at a point l. In the computations, we arbitrarily choose l such that $l = N/4$.

7) ω is adapted based on the value of Dif. If $\text{Dif} < 100$, the value is increased by a factor of 10 and is further increased for each order of magnitude that Dif drops until a prescribed maximum value $\omega_{\max}$ is reached. In these computations $\omega_{\max} = 0.1$.

8) If $\text{Dif} > \epsilon$ where ϵ is a small number, return to step 2. The code used here has

$\epsilon = 10^{-5}$.

References

Abercrombie, R., Earthquake source scaling relationships for -1 to 5 M_L using seismograms recorded at 2.5 km depth, *J. Geophys. Res.*, *100*, 24,015–24,036, 1995.

Abramowitz, M., and I. Stegun, *Handbook of mathematical functions, with formulas, graphs, and mathematical tables*, Dover, New York, 1965.

Carmichael, R. S., ed., *Handbook of Physical Properties of Rocks*, vol. 2, CRC, Boca Raton, FL, 1982.

Davis, S. N., and R. J. DeWiest, *Hydrogeology*, Wiley, New York,, 1966.

Dowson, D., and G. Higginson, *Elasto-Hydrodynamic Lubrication*, Pergamon, 1977.

Fletcher, J., and P. Spudich, Rupture characterisitics of the three M $\sim$ 4.7 (1992–1994) Parkfield earthquake, *J. Geophys. Res.*, *103*, 835–854, 1998.

Hamrock, B. J., *Fundamentals of Fluid Film Lubrication*, McGraw-Hill, 1994.

Hubbert, M. K., and W. W. Rubey, Role of fluid pressure in mechanics of overthrust faulting, *Bull. Geol. Soc. Amer.*, *70*, 115–166, 1959.

Husseini, M. I., Energy balance for motion along a fault, *Geophys. J. R. Astr. Soc.*, *49*, 699–714, 1977.

Jacobson, B., *Rheology and Elastohydrodynamics Lubrication*, Elsevier, Amsterdam, 1991.

Jeffreys, H., On the mechanics of faulting, *Geol. Mag.*, *79*, 291, 1942.

Jiang, X., D. Hua, H. Cheng, X. Ai, and S. C. Lee, A mixed elastohydrodynamic lubrication model with asperity contact, *J. of Tribology*, pp. 481–491, 1999.

Kanamori, H., and D. L. Anderson, Theoretical basis of some empirical relations in seismology, *Bull. Seism. Soc. Amer.*, *65*, 1073–1095, 1975.

Kanamori, H., and T. Heaton, *GeoComplexity and the Physics of Earthquakes*, chap. Microscopic and Macroscopic Physics of Earthquakes, pp. 147–183, AGU, Washington, D.C., 2000.

Kaneko, S., H. Takabatake, and K. Ito, Numerical analysis of static characteristics at start of operation in porous journal bearings with sealed ends, *J. of Tribology*, *121*, 62–68, 1999.

Kushiro, I., *Physics of Magmatic Processes*, chap. Viscosity, Density, and Structure of slicate melts at high pressures, and their petrological applications, pp. 93–117, Princeton Univ., Princeton, NJ, 1980.

Lachenbruch, A. H., Frictional heating, fluid pressure and the resistance to fault motion, *J. Geophys. Res.*, *85*, 6097–6112, 1980.

Lachenbruch, A. H., and J. H. Sass, Heat flow and energetics of the San Andreas fault zone, *J. Geophys. Res.*, *85*, 6185–6222, 1980.

Lay, T., and T. Wallace, *Modern Global Seismology*, Academic Press, San Diego, 1995.

Leal, L., *Laminar Flow and Convective Transport Processes*, Butterworth-Heinemann, Boston, 1992.

Ma, K.-F., C.-T. Lee, and Y.-B. Tsai, The Chi-Chi, Taiwan earthquake: Large surface displacements on an inland thrust fault, *EOS*, *80*, 605–611, 2000.

Ma, K.-F., T.-R. A. Song, J. Mori, and E. E. Brodsky, Direct observations of fault dynamics for the 1999 Chi-Chi Taiwan earthquake (M_w 7.7), in prep.

Major, J. J., and T. C. Pierson, Debris flow rheology: Experimental analysis of fine-grained slurries, *Water Resourcs Res.*, *28*, 841–857, 1992.

Mase, C. W., and L. Smith, Effects of frictional heating on the thermal, hydrologic, and mechanical response of a fault, *J. Geophys. Res.*, *92*(B7), 6249–6272, 1987.

Mott, N. F., Fracture of metals: Some theoretical considerations, *Engineering*, *165*, 16, 1948.

Okamura, H., A contribution to the numerial analysis of isothermal elastohydrodynamic lubrication, in *Tribology of Reciprocating Engines*, edited by D. Dowson, pp. 313–320, 1982.

Otsuki, K., Thermal pressurization, fluidization and melting of fault gouge recorded in the rock from Nojima seismic fault, *EOS*, *80*, 727, 1999.

Persson, B. N. J., *Sliding Friction: Physical Principles and Applications*, Springer-Verlag, New York, 1998.

Power, W. L., and T. E. Tullis, Euclidean and fractal models for the description of rock surface roughness, *J. Geophys. Res.*, *93*, 415–424, 1991.

Prejean, S. G., and W. L. Ellsworth, Observations of earthquake source parameters and attenuation at 2 km depth in the Long Valley Caldera, Eastern California, *Bull. Seism. Soc. Am.*, submitted.

Reynolds, O., On the theory of lubrication and its application to Mr. Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil, *Phil. Trans. Roy. Soc.*, *177*, 157–234, 1886.

Rumble, D., Water circulation in metamorphism, *J. Geophys. Res.*, *99*, 15499–15502, 1994.

Sato, T., and T. Hirasawa, Body wave spectra from propagating shear cracks, *J. Phys. Earth*, *21*, 415–431, 1973.

Scholz, C. H., *The Mechanics of Earthquakes and Faulting*, Cambridge, 1990.

Sengers, J., and J. Watson, Improved international formulations for the viscosity and thermal-conductivity of water substance, *J. of Phys. Chem. Ref. Data*, *15*, 1291–1314, 1986.

Sibson, R., Interactions between temperature and pore-fluid pressure during earthquake faulting and a mechanism for partial or total stress relief, *Nature Phys. Sci.*, *243*, 66–68, 1973.

Sibson, R., Thickness of the seismogenic slip zone: Constraints from field geology, *EOS*, *80*, 1999.

Sommerfeld, A., *Mechanics of Deformable Bodies*, Academic Press, New York, 1950.

Spikes, H., Mixed lubrication – an overview, *Lubrication Science*, *9*, 221–253, 1997.

Spray, J. G., Viscosity determinations of some frictionally generated silicate melts: Implications for fault zone rheology at high strain rates, *J. Geophys. Res.*, *98*, 8053–8068, 1993.

Szeri, A. Z., *Fluid film Lubrication: Theory & Design*, Cambridge Univ., Cambridge, UK, 1998.

Timoshenko, S., and J. Goodier, *Theory of Elasticity*, McGraw-Hill, NewYork, 1970.

Wald, D. J., and T. H. Heaton, Spatial and temporal distribution of slip for the 1992 Landers, California, Earthquake, *Bull. Seism. Soc. Amer.*, *84*, 668–691, 1994.

Wald, D. J., H. Heaton, Thomas, and K. Hudnut, The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data, *Bull. Seism. Soc. Amer.*, *86*, S49–S70, 1996.

Part II

Triggering by Seismic Waves at Regional Distance

Preface

Regional-scale triggering of seismicity and eruptions by large earthquakes has been reported anecdotally for many years, but only recently has the phenomenon been robustly and convincingly demonstrated. Modern, dense sense networks recorded a sharp increase in seismicity following the 1992 Landers earthquake [*Hill et al.*, 1993]. The triggered sites were up to 1200 km away. *Hill et al.* [1993] was the first modern study of triggered seismicity and convinced even some of the most sceptical researchers that farfield triggering can and does occur [*Ruff*, 1993].

One of the surprises of the Landers observation is that triggered activity is associated with geothermal and magmatic areas. Triggered sites in the western United States included Long Valley, the Geysers and Yellowstone. If this apparent correlation is confirmed by studies of triggering cases in other regions, it implies that fluids are an important component of the triggering process. Chapter 3 presents further evidence that regional triggering occurs. The 1999 Izmit earthquake is shown to have triggered seismicity in Greece. The region is under extension and there is some evidence that the triggered sites may be geothermal.

Eruptions that appear to be triggered by large, distance earthquakes have also been reported anecdotally. Recently *Linde and Sacks* [1998] showed that eruptions occur more often within a day of large earthquakes than would be expected of a random sampling. In light of our new understanding that earthquakes can have significant effects hundreds of kilometers from the epicenter, triggered eruptive mechanisms deserve renewed attention. In chapter 4 we evaluate the proposal that rectified diffusion, a process used in engineering applications, allows transient seismic waves to increase the pressure in magmatic systems. A similar mechanism has been proposed to explain triggered seismicity [*Sturtevant et al.*, 1996].

References

Hill, D., et al., Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, *Science*, *260*(5114), 1617–1623, 1993.

Linde, A., and I. Sacks, Triggering of volcanic eruptions, *Nature*, *395*, 888–890, 1998.

Ruff, L., Earthquakes - learning from the whispers, *Nature*, *364*, 576–577, 1993.

Sturtevant, B., H. Kanamori, and E. E. Brodsky, Seismic triggering by rectified diffusion in geothermal systems, *J. Geophys. Res.*, *101*, 25,269–25,282, 1996.

Chapter 3 A New Observation of Dynamically Triggered Regional Seismicity: Earthquakes in Greece Following the August 1999 Izmit, Turkey, Earthquake

Abstract

The M_w=7.4 Izmit, Turkey, earthquake triggered widespread regional seismicity in Greece over a study region extending from 400 km to nearly 1000 km away from the epicenter. Small events began immediately after the passage of the mainshock surface waves suggesting that the transient stresses of the seismic waves were the trigger. The increase in cataloged earthquakes in ordinary continental crust is a new observation and is statistically significant at the 95% level. Unlike the previous example of distant triggering during the Landers earthquake, the activated seismicity occurred entirely in non-volcanic areas. The Greek sites were triggered by waves with amplitudes at least a factor of 3 lower than the amplitude of the observed triggering threshold for Imperial Valley. We speculate that dynamic triggering on a regional-scale results in countrywide episodes of increased seismicity, or "superswarms," in regions with low triggering thresholds such as Greece.

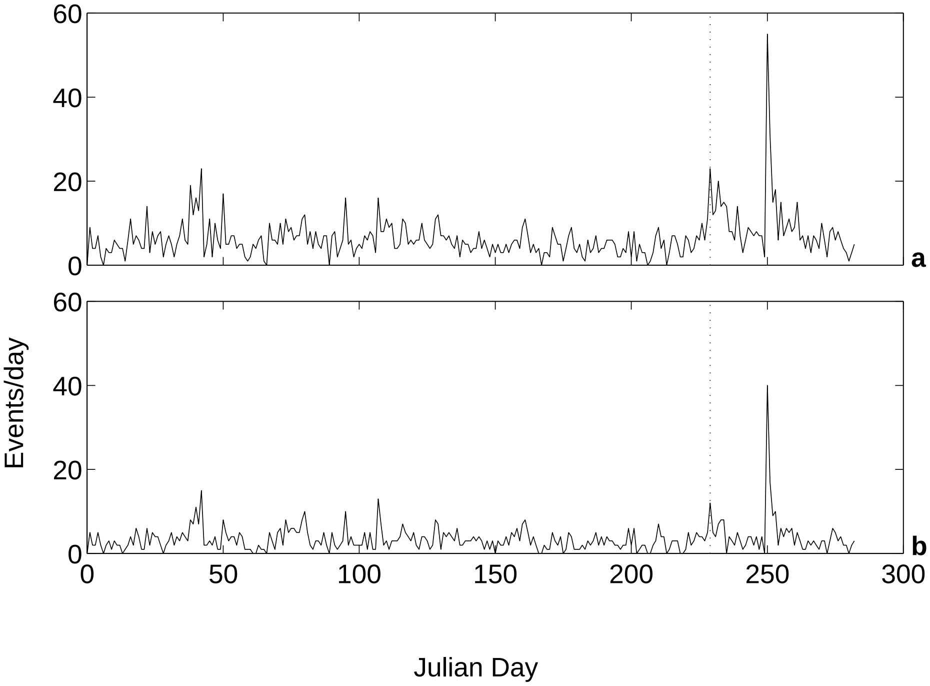

Figure 3.1 Number of earthquakes per day recorded by Thessaloniki network west of longitude 25°. (a) Events of $M \geq 2$. (b) Events of $M \geq 3.5$. The dotted line indicates the date of the Izmit earthquake. The peak on day 250 is caused by the Athens M_w=5.8 event and its aftershocks.

Observation

The M_w=7.4 Izmit earthquake on August 17, 1999, at 00:01:39.80 UT was followed immediately by small earthquakes occurring throughout much of continental Greece. The number of earthquakes per day recorded by the network of the Department of Geophysics of the University of Thessaloniki from January 1, 1999, through October 9, 1999, are shown in Figure 3.1. A peak is visible on the day of the Izmit event (Julian day 229). The seismicity discussed in this figure and throughout this study includes only the events west of longitude 25°, i.e., at least 400 km from the Izmit epicenter at 40.702°N 29.987°E. A cursory inspection of the daily seismicity maps before and after day 229 strongly suggests a link between the activity in Greece and the Turkish event (Figure 3.2).

The triggering of events from the Izmit mainshock occurred immediately after the

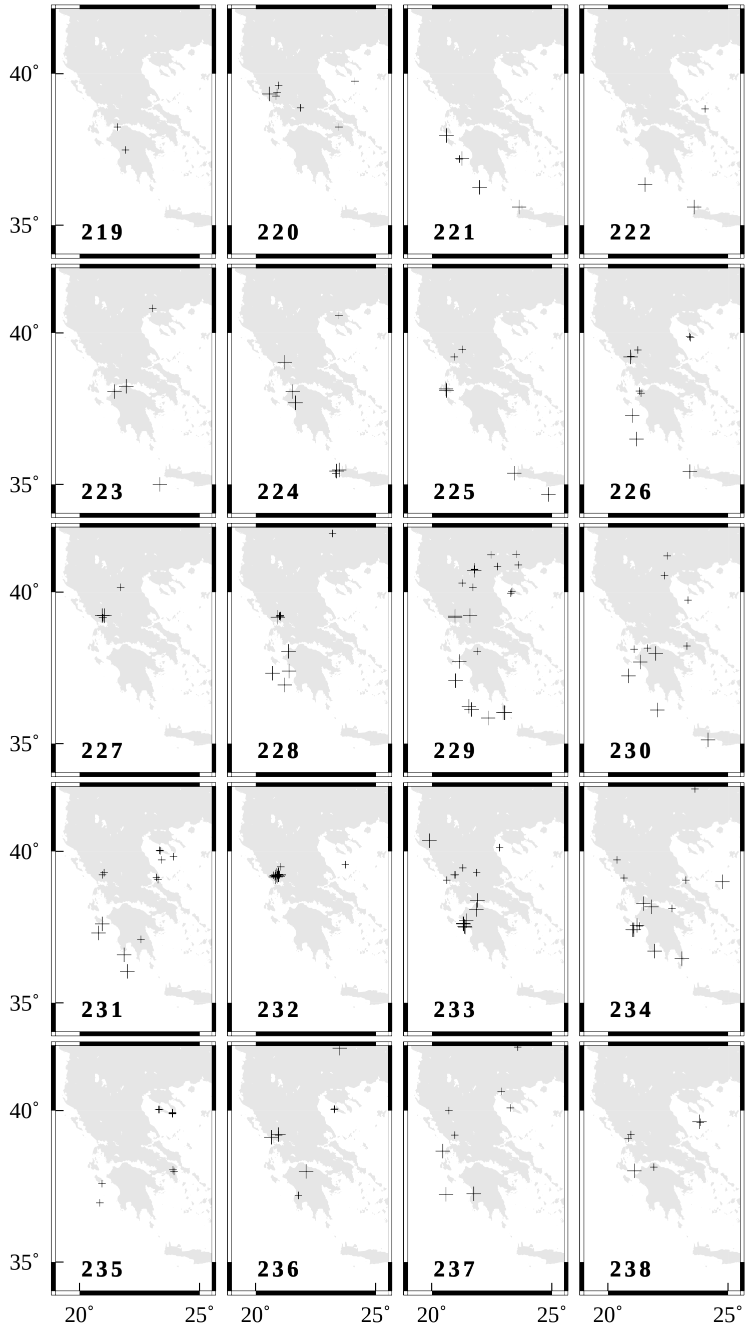

Figure 3.2 Map of events recorded by the Thessaloniki network west of 25° during the days surrounding the Izmit earthquake (Julian day 229). The large crosses are events with $M \geq 3.5$ and the small crosses indicate $3.5 > M \geq 2$.

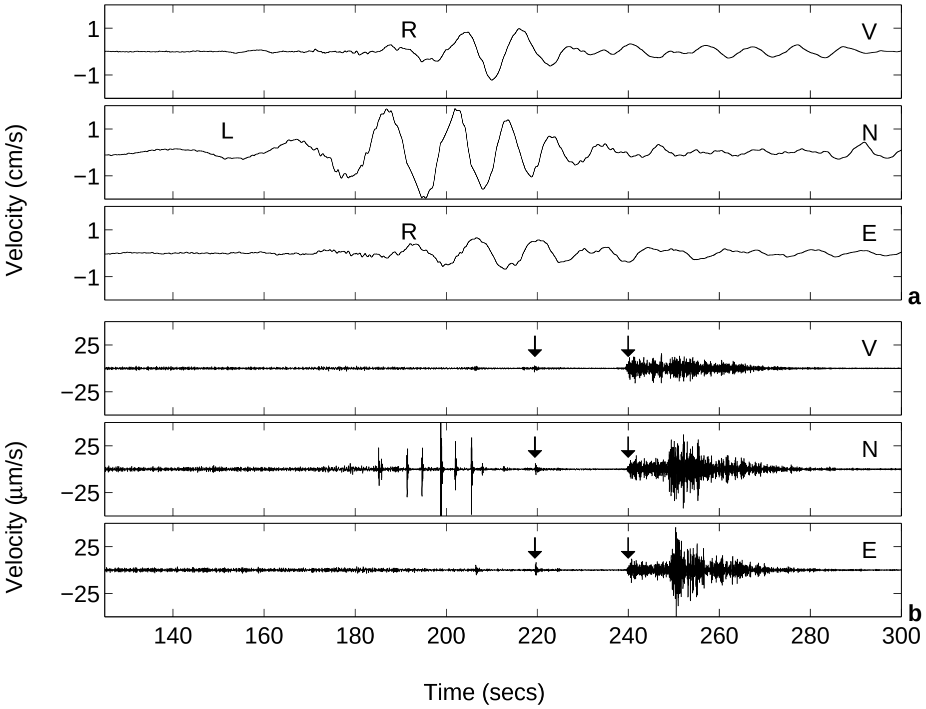

Figure 3.3 The waveforms of the mainshock and local events in Greece occurring immediately after it. (a) Records from a temporary deployment of a Guralp CMG40-T at 40.5952°N 23.0023°E (station 7905). The three components are as labeled. (b) The same records band-passed 5–20 Hz. Note the difference in scales. The two events marked by arrows are located at 40.580°N 22.860°E and 39.930°N 23.367°E with origin times of 00:05:13.4 UT and 00:05:26.1 UT and magnitudes $M_L \approx 1$ and 3.8, respectively. The glitches on the north component result from band-passing the clipped signal. The origin time of the plot is the origin time of the Izmit mainshock. The Love (L) and Rayleigh (R) arrivals are labeled.

passage of the surface waves. Figure 3.3 shows several local, high-frequency events that occurred in Greece directly after the surface waves. Although it is possible that the occurrence of these local earthquakes was coincidental, the timing strongly suggests that the large amplitude dynamic strain of the surface waves is responsible for triggering regional seismicity. The local events shown here are not in standard catalogs since the surface waves obscure their arrivals. This masking of early events by the mainshock coda results in an apparent delay of approximately a half hour before the onset of triggered seismicity in the Thessaloniki catalog.

This observation of regionally triggered seismicity is significant for the following reasons: (1) It is the first fully documented observation of regional triggering over a large area by seismic waves in non-volcanic continental crust. (2) The threshold for triggering is at least a factor of 3 lower than in volcanic regions of California. (3) The observation demonstrates interactions between earthquakes over large distances and suggests that regions with low triggering thresholds are prone to episodes of widespread increased seismicity.

Catalog

We use the catalog of the Thessaloniki network for this study. The network consists of 16 Teledyne S-13 stations with natural frequencies of 1 Hz located in northern Greece (Figure 3.5). All stations were operational throughout 1999 except station LKD which was activated on July 10, 1999, and located at 38.707°N 20.651°E. There was no change in the automated cataloging procedures after the Izmit event; therefore, we can eliminate extraordinary alertness of the seismological staff following a large event as a cause of the apparently elevated seismicity. The Thessaloniki catalog exhibits the standard Gutenberg-Richter magnitude-frequency relationship above M=3.5 and is depleted in events below this magnitude. Following standard practices, we take M=3.5 as the threshold for completeness of the catalog. As can be seen from Figures 3.1 and 3.2, the peak in seismicity post-Izmit is still evident even if we limit ourselves to the formally complete section of the catalog. Including the smaller events in Figures 3.1a and 3.2 provides a larger sampling and makes the phenomenon even clearer.

Statistical significance

We fit a log-normal probability distribution function to the catalog for the first 228 days of 1999 in order to determine the statistical significance of the peak on day 229. We then used the best-fitting log mean and log variance for the catalog of $M \geq 3.5$ events to calculate the probability that the 12 events occurring on day 229 was a coincidence unrelated to the Izmit earthquake. With this distribution the probability that 12 or more events would occur on a given day is $\sim 3\%$. The statistical calculation was repeated for larger catalogs including all events from 1988 to 1998 or all events with $M \geq 2$. Even in the least favorable case, the hypothesis that the increase in seismicity on day 229 was coincidental can be rejected at the 95% level. Furthermore, the other days in 1999 with a similar number of events were during locally confined (<250 km^2) earthquake sequences. The activity on day 229 was unique in the 1999 catalog in that it had a large number of events spatially dispersed over 4×10^5 km^2 (Figure 3.4).

Spatial distribution

It has been previously suggested that dynamically triggered seismicity occurs preferentially in geothermal and magmatic areas [*Hill et al.*, 1993]. Continental Greece has no recent magmatism, but there is a possible correlation between triggered seismicity and geothermal areas. *Anderson et al.* [1994] has suggested that such apparent correlations are artifacts of the fact that an increase in events is most easily measured in the regions which are most seismically active. Such active areas are commonly geothermal. To distinguish between these possibilities we have relocated the events listed in the Thessaloniki catalog using the combined phase picks of the Thessaloniki network and the Geodynamical Institute of the National Observatory of Athens. Location

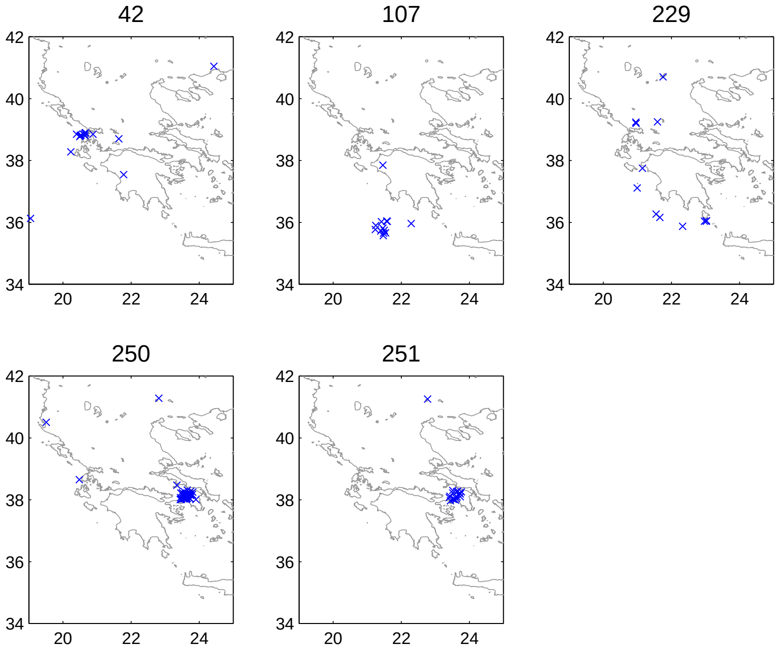

Figure 3.4 Map of daily seismicity ($M \geq 3.5$) for all days in the Thessaloniki catalog January 1 – October 9 1999 with a 12 or more $M \geq 3.5$ events. Julian days are as marked. All days except 229 have the majority of the seismicity in a tight cluster.

errors for the combined network were estimated by comparing the network locations of the September 7, 1999, M_w=5.9 Athens aftershock sequence with high-quality locations obtained from a local, temporary array. The mean difference between the network and temporary array locations is 8 km horizontally and 5 km in depth.

The map in Figure 3.5 shows the relocated triggered events and the level of background seismicity. The Council of National Seismic System (CNSS) composite catalog of events with $M \geq 4$ during the interval 1961–1998 is used as a measure of the normal regional seismicity. This catalog has relatively uniform spatial coverage for nearly 40 years and is complete according to a standard test of the magnitude-frequency distribution. The geographical distribution is dominated by a few large mainshock-aftershock sequences. The difference between active (colored) and inactive (uncolored) areas is therefore interpreted to be more significant than the differences in levels of activity among the seismic regions.

We do not attempt a systematic statistical study of the spatial correlation of the

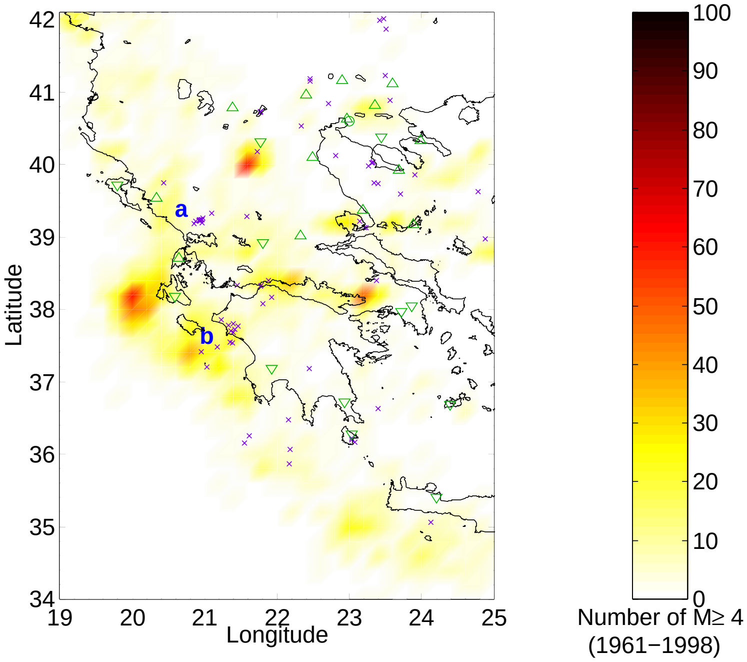

Figure 3.5 Plot of relocated events with $M \geq 2$ from 8/17/99–8/22/99 inclusive (crosses) and background activity (colorbar). The background seismicity is binned into $0.2^o \times 0.2^o$ cells and then smoothed by linear interpolation between cells. Seismic network stations are plotted in green. Thessaloniki stations are upward pointing triangles, Athens stations are downward pointing triangles and station 7905 is a circle at 40.5952^oN 23.0023^oE. Groups of events are labeled (a) Arta cluster (b) Pirgos cluster.

triggered events with either geological features or background activity since the data are too sparsely distributed. We focus instead on the two most striking groups of the sequence: (a) the cluster of events near Arta and (b) the cluster near Pirgos. Both areas have grabens and documented thermal springs [*Waring*, 1965]. The springs indicate geothermal activity in that hot, aqueous fluids are present at shallow depths, although the heat flow is not necessarily elevated relative to the surrounding region. The Arta cluster occurred in a region which generally has a relatively low level of seismicity as shown by the white background in Figure 3.5. Contrary to the hypothesis of *Anderson et al.* [1994], the largest number of observed triggered events are not in the most seismically active areas. Triggered activity outside the clusters may also be located at geothermal sites since hot springs occur over much of Greece [*Waring*, 1965]. All triggered events are confined to crustal depths, even though ∼20% of the 1999 Thessaloniki catalog events are below 33 km.

Comparison with Southern California

We compare the Izmit triggered seismicity to the two other well-documented examples of regionally triggered seismicity in order to establish the relative sensitivity of triggered regions. The June 28, 1992, M_w=7.3 Landers earthquake [*Hill et al.*, 1993] and the October 16, 1999, M_w=7.1 Hector Mine earthquake both triggered widespread microseismicity in southern California [*U.S. Geol. Surv., S. Calif. Earthq. Cent. and Calif. Div. Mines Geol.*, in press]. Like Izmit, both events were large strike-slip earthquakes and they triggered seismicity in areas under regional extension with geothermal activity. Since it is established that both the Landers and Hector Mine earthquakes had the ability to trigger distant events, we infer that any area that was triggered during one of the events but not the other experienced shaking above its local triggering threshold during only one of these events. Imperial Valley, California, was triggered during Hector Mine and not Landers. Therefore, the recorded ground motion during Hector Mine provides an upper bound to the triggering threshold in Imperial Valley and the record of Landers provides a lower bound. It is assumed that there was no change in the triggering threshold during the seven years between Landers and Hector Mine.

The strength of the triggering waves can be measured by either the amplitude of the transient stress, which scales as the particle velocity, or by the energy density delivered by the waves. Table 3.1 lists both of these metrics. It shows that triggering in Greece occurred at amplitudes and energies lower than Imperial Valley. The record in Figure 3.3 measures shaking within 20 km of at least one $M_L \approx 1$ event. Four nearby stations situated on a variety of rock types show similar records which suggests that local site response effects are negligible for the long-period waves considered here. The record is a good measure of the shaking at the triggered site and yields the values in Table 3.1. Since the site is closer to the Izmit mainshock than most of the other Greek triggered sites, the values measured here are upper bounds on the triggering threshold for susceptible regions in Greece. Whether amplitude or energy is taken as a measure of the strength of the triggering wave, the threshold is more than a factor of three lower at the triggered sites in Greece than in Imperial Valley.

Triggering mechanisms

Although static stress changes may be an important trigger for nearfield aftershocks, the static stress changes at more than a fault length from the source are negligible even compared to tidal stresses. The observation of events immediately after the largest amplitude shaking (Figure 3.3) is consistent with the dynamic stress triggering the earthquakes. A physical mechanism is required to transform the transient stresses of the seismic waves into sustained stresses on the fault capable of producing an earthquake hours or days later and a number of possible mechanisms have been suggested [*Hill et al.*, 1993; *Linde et al.*, 1994; *Sturtevant et al.*, 1996; *Gomberg et al.*, 1998].

Three mechanisms have received detailed treatment. *Linde et al.* [1994] suggest that bubbles in magmatic systems are shaken loose by the seismic waves. As the bubbles rise, the advective overpressure stresses the system and causes earthquakes.

Table 3.1. Observed Amplitudes and Energies

Event	Station	Δ km	Amplitude cm/s (MPa)	En. Density J/m^2
Landers	SSW[1]	139	6.6 (0.60)	3.5×10^5
Hector Mine	SSW	158	9.0 (0.82)	1.3×10^6
Izmit	7905	589	1.9 (0.18)	7.8×10^4

Distance Δ is measured from the epicenters to the stations. Amplitude is reported in terms of both velocity and dynamic stress. The velocity amplitude is half the maximum peak-peak value of the horizontal velocity. The stress amplitude is the velocity amplitude multiplied by μ/β where μ is the shear modulus (3×10^{10}Pa) and β is the shear velocity (3300 m/s). Energy density is the integral over time of the velocity squared multiplied by β and the rock density (2750 kg/m^3) [*Kanamori et al.*, 1993]. Station SSW is in Imperial Valley. Station 7905 in Greece is described in Figure 3.3.

[1] SSW was not operational in 1992. The amplitude and energy density here are calculated by applying a series of corrections to the velocity traces of the nearest station (PFO, Δ=63 km). A Hector Mine aftershock (10/16/99 20:13:37.0 M_w=4.7) recorded on-scale at both PFO and SSW is used to determine the correction. The raw amplitude ratio between PFO and SSW for the aftershock was corrected for the theoretical radiation pattern of the event using the standard theory of *Kanamori and Stewart* [1976] and the focal mechanism produced by Lupei Zhu. The resulting ratio is interpreted as the relative site amplification. The site amplification is applied to the Landers records at PFO to find the shaking at SSW. In addtion, an empirical correction for the Landers radiation pattern and directivity [*Kanamori et al.*, 1992] is applied to the Landers records.

This theory was formulated for magmatic systems rather than the geothermal ones observed here. The theory cannot be readily adapted to the geothermal case since it requires unobstructed bubble rise which is not possible in a porous medium. Therefore, the mechanism is not applicable to triggering in continental Greece. *Sturtevant et al.* [1996] suggest that the seismic waves pressurize bubble-filled regions of the geothermal system via rectified diffusion. This mechanism is physically plausible under certain, limited conditions that will be discussed in the next chapter. *Gomberg et al.* [1998] suggest that transient strains of the seismic waves temporarily increase the sliding velocity in regional faults. The velocity increase is enough to reduce the friction as prediced by the rate-state formulation of *Dieterich* [1994] and this process results in advancing a number of near-critical fault patches to a failure state. The *Gomberg et al.* [1998] theory is appealing because it intrinsically addresses the variable strength criteria. It has the disadvantage of not addressing the observation that geothermal fluids appear to be a prerequisite for triggering.

It is possibile that the tensional stress regime rather than the presence of fluids distinguishes areas prone to triggering. Initiating slip on optimally oriented normal faults requires lower amplitude shear stresses to overcome the static friction than on optimally oriented thrust or strike-slip faults [*Sibson*, 1974]. However, the mechansims of triggered events in Long Valley post-Landers include both normal and strike-slip events [*Hill et al.*, 1995]. If the threshold shear stress was the only factor involved and the shear stresses were enough to overcome the static friction on strike-slip faults, then triggered seismicity on the nearby Garlock fault should have followed Landers. No such events were observed.

A previously undiscussed mechanism for dynamic triggering is subcritical crack growth accelerated by seismic waves. At crack tips in wet rocks, chemical reactions are accelerated by the high stresses and cracks slowly grow [*Das and Scholz*, 1981]. Large cracks grow faster than smaller ones due to the increased stress intensity at the tip. If the cracks are in a near-critical state prior to the seismic shaking, the large amplitude seismic waves temporarily increase the rate of stress corrosion. This transient stress increases the average size of the crack population and accelerates critical failure. As stress corrosion is most effective in hot, wet rocks, this model predicts that geothermal areas are likely to have a high occurrence of triggered events.

Regional superswarms

Observable interactions between distant earthquakes as shown in this study suggest that seismicity can be coupled over broad regions. We call such regional-scale episodes of elevated seismicity “superswarms.” Superswarms can last up to a few months and are distinguished from common mainshock-aftershock sequences by the large areas involved. Candidates for superswarms include the October 1994 to January 1995 sequence of large (M>7.0) earthquakes in Japan and the September 1999 Athens event following the August 1999 Izmit earthquake. Areas prone to superswarms have low triggering thresholds, as was observed for Greece. Such regions can be identified globally by using modern broadband instruments to record the amplitude and energy of triggering waves. A systematic evaluation of regional sensitivity using the methods shown in Figure 3.3 and Table 3.1 would be a logical extension of this work and a valuable test of the hypothesis.

A related observation unique to Greece is the apparent occurrence of electrical signals during periods of heightened seismicity. We speculate that the observation that large events trigger widespread seismicity in the Mediterranean region suggests an explanation for both the occurrence of electrical signals during seismically active periods as well as the observed clustering of large events in the region. The controversial VAN (Varotsos-Alexopoulos-Nomicos) method has been used in Greece to study the apparent correlation between observed electrical signals and large earthquakes [*Varotsos et al.*, 1993]. Other authors have noted that the earthquakes which have seismoelectric signals according to VAN are commonly preceded by foreshocks [*Sudo*, 1996]. The 1995 group of events studied by VAN spanned all of continental Greece [*Varotsos et al.*, 1996].

We suggest the following scenario linking the seismoelectric observations and superswarms. The seismic waves from a large event cause the pressurization of geothermal systems throughout the region either by rectified diffusion [*Sturtevant et al.*, 1996] or other mechanisms resulting from the interaction of the seismic waves with the multiphase fluid systems. The pressurization has two effects. (1) Earthquakes are triggered in the pressurized areas. (2) The pressurized fluid flows in accordance with Darcy's law. The streaming potential associated with these flow events may be the signal observed by the VAN system. The scenario would explain why electrical signals appear to be related to periods of high seismicity. Both the superswarms and the electrical signals are symptomatic of the same process of regional coupling between earthquakes and pore fluids by seismic waves. Further studies documenting dynamic triggering and associated superswarms will test the applicability of the proposed scenario.

References

Anderson, J., J. Brune, J. Louie, Y. Zeng, M. Savage, G. Yu, Q. Chen, and D. DePolo, Seismicity in the western Great-Basin apparently triggered by the Landers, California, earthquake, 28 June 1992, *Bull. Seism. Soc. Am.*, *84*(3), 863–891, 1994.

Das, S., and C. H. Scholz, Theory of time-dependent rupture in the Earth, *J. Geophys. Res.*, *86*, 6039–6051, 1981.

Dieterich, J., A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, *99*, 2601–2618, 1994.

Gomberg, J., N. Beeler, M. Blanpied, and P. Bodin, Earthquake triggering by transient and static deformations, *J. Geophys. Res.*, *103*(B10), 24411–24426, 1998.

Hill, D., et al., Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, *Science*, *260*(5114), 1617–1623, 1993.

Hill, D. P., M. J. S. Johnston, and J. O.Langbein, Response of Long Valley caldera to the M_w=7.3 Landers, California, earthquake, *Journ. Geophys. Res.*, *100*, 12,985–13,005, 1995.

Kanamori, H., and G. S. Stewart, Mode of the strain release along the Gibbs fracture zone, Mid-Atlantic Ridge, *Phys. of the Earth and Planetary Interiors*, *11*, 312–332, 1976.

Kanamori, H., H.-K. Thio, D. Dreger, and E. Haukson, Initial investigation of the Landers, California, earthquake of 28 June 1992 useing TERRAscope, *Geophys. Res. Let.*, *19*, 2267–2270, 1992.

Kanamori, H., J. Mori, E. Hauksson, T. H. Heaton, L. K. Hutton, and L. M. . Jones, Determinations of earthquake energy release and M_L using Terrascope, *Bull. Seism. Soc. Am.*, *83*, 330–346, 1993.

Linde, A., I. Sacks, M. Johnston, D. Hill, and R. Bilham, Increased pressure from rising bubbles as a mechanism for remotely triggered seismicity, *Nature*, *371*(6496), 408–410, 1994.

Sibson, R., Frictional constraints on thrust, wrench and normal faults, *Nature*, *249*, 542–544, 1974.

Sturtevant, B., H. Kanamori, and E. E. Brodsky, Seismic triggering by rectified diffusion in geothermal systems, *J. Geophys. Res.*, *101*, 25,269–25,282, 1996.

Sudo, K., *A critical review of VAN*, chap. Foreshocks preceding VAN signals, pp. 239–243, World Scientific, Singapore, 1996.

U.S. Geol. Surv., S. Calif. Earthq. Cent. and Calif. Div. Mines Geol., Preliminary report on the 10/16/1999 M7.1 Hector Mine, California earthquake, in press.

Varotsos, P., K. Alexopoulos, and M. Lazaridou, Latest aspects of earthquake prediction in Greece based on seismic electric signals .2., *Tectonophysics*, *224*(1-3), 1–37, 1993.

Varotsos, P., M. Lazaridou, K. Eftaxias, G. Antonopoulos, J. Makris, and J. Kopanas, *A critical review of VAN*, chap. Short Term Earthquake Prediction in Greece by seismic electric signals, pp. 29–76, World Scientific, Singapore, 1996.

Waring, G. A., Thermal springs of the United States and other countries of the world—a summary, *U.S. Geol. Surv. Prof. Pap.*, *492*, 383, 1965.

Chapter 4 An Evaluation of Rectified Diffusion as a Means for Triggering Volcanic Eruptions

Abstract

Rectified diffusion is a mechanism by which a strain wave can rapidly pump volatiles into a bubble and therefore increase the pressure in a closed system. The dynamic strain of distant regional tectonic earthquakes can be translated to static strain inside a magma chamber via this process. We formulate a theory appropriate to the conditions of a magma chamber and calculate the increased pressure using realistic physical parameters. For a basaltic system initially at 130 MPa pressure, the excess pressure from rectified diffusion is between 0.001 and 0.02 MPa for a regional $M \geq 8$ earthquake. Strong constraints on the parameters including porosity, bubble size, velocity structure and permeability must be imposed in order for rectified diffusion to be effective. The pressure from rectified diffusion is often comparable to the static stress caused by deformation for documented cases of triggered eruptions.

Introduction

Earthquakes are often used empirically to forecast volcanic eruptions. However, the precise relationship between seismicity and magmatic activity remains enigmatic. In some cases, e.g., swarms, earthquakes are taken to be indicative of magma movement and therefore symptomatic of volcanic unrest. In other cases the relationship is apparently causal. There are a few documented cases of distant regional earthquakes triggering eruptions. In 1835, Darwin observed four volcanoes erupting after a large earthquake (M=8.5) on the Chilean coast [*Darwin*, 1896]. The only historical eruptions of both the Tao-Rusyr caldera and the Karpinsky group volcanoes occurred within days of a M=8.3 earthquake [*Kimura*, 1978]. *Newhall and Dzurisin* [1988] document 50 cases of large regional earthquakes followed by caldera unrest. *Linde and Sacks* [1998] found a statistically significant correlation between large eruptions and earthquakes occurring within a day of each other by searching the entire historical catalog. In this study we limit ourselves to 11 of the best documented eruptions immediately following distant regional earthquakes. The eruptions we consider, shown in Figure 4.1, were triggered by large tectonic events over 100 km away. Previous work has attempted to explain triggered eruptions by computing the static stress change from elastic deformation [*Yamashina and Nakamura*, 1978]. The stress change thus computed is as low as 10^{-4} MPa in some cases. These stresses are comparable to tidal stresses and therefore it is difficult to justify triggering a volcanic eruption by this means. It has been suggested that rectified diffusion is an alternative mechanism to relate large, regional earthquakes to eruptions [*Sturtevant et al.*, 1996; *Brodsky et al.*, 1998]. Here we elaborate on the applicability of rectified diffusion to volcanic systems and quantitatively constrain the conditions under which this mechanism may be applicable.

Rectified diffusion of dissolved vapor into preexisting bubbles in a magma body is a mechanism that uses the dynamic strain from distant regional earthquakes to trigger activity. Such bubbles exist in many natural systems as documented from melt inclusions [*Lowenstern*, 1995], volatile contents that exceed saturation at moderate

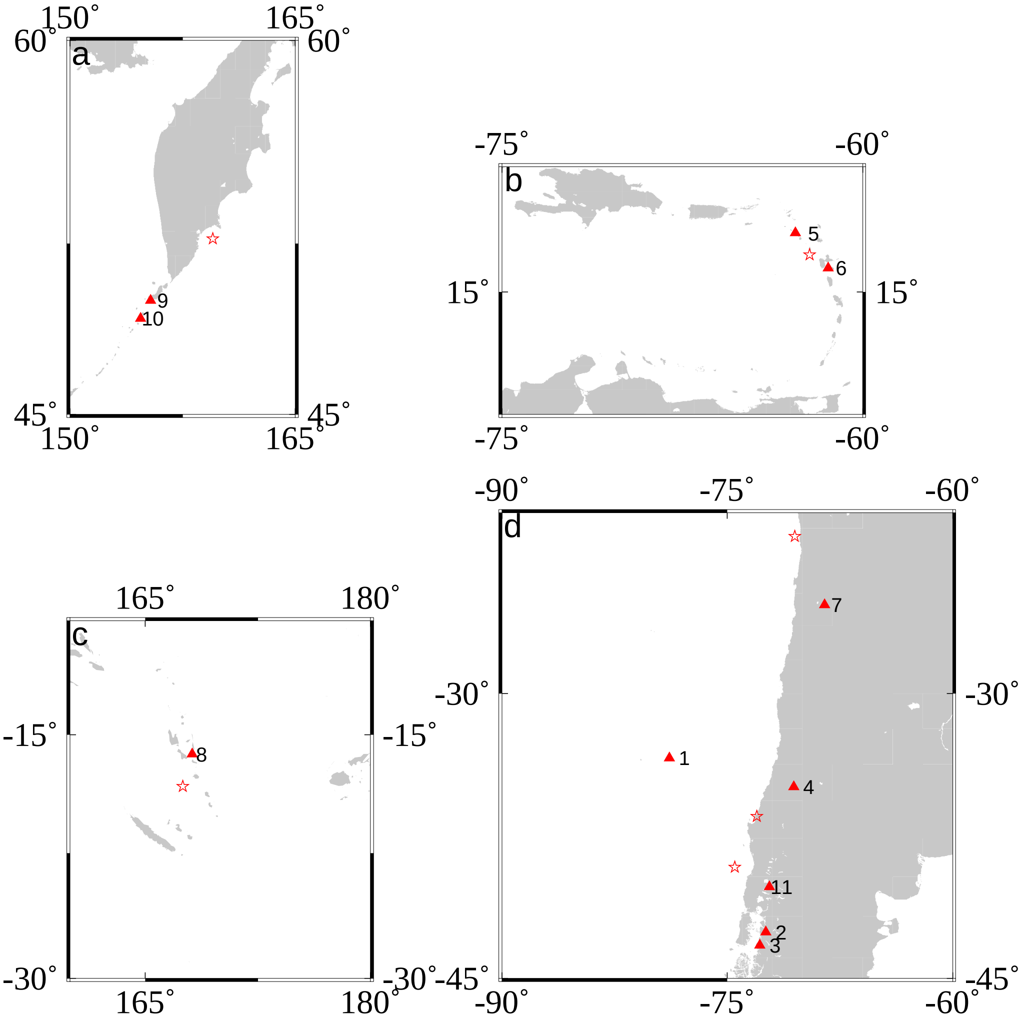

Figure 4.1 Location of volcanoes and earthquakes studied in this work. The regions mapped are (a) Kamchatka, (b) the Caribbean, (c) the Southwest Pacific and (d) Chile. Triangles represent volcanoes and stars represent earthquakes. The numbers correspond to the following eruptions detailed in Table 4.3: 1, 1835 Robinson Crusoe; 2, 1835 Minchinmavida; 3, 1835 Cerro Yanteles; 4, 1835 Peteroa; 5, 1843 Liamuiga; 6, 1843 unnamed; 7, 1877 Llullaillaco; 8, 1950 Ambrym; 9, 1952 Karpinsky group; 10, 1952 Tao-Rusyr caldera; 11, 1960 Puyahe.

depths [*Johnson et al.*, 1994], and evidence of coexisting gas phases in magma chambers [*Lambert et al.*, 1985]. When subjected to seismic waves, these bubbles expand and contract. Assuming the bubbles are originally near equilibrium with the vapor dissolved in the melt, when the bubbles contract, the vapor is oversaturated inside the bubble and diffuses out to the melt (Figure 4.2). When the bubbles expand, the vapor is undersaturated inside and diffuses in from the melt. Since the surface area of the expanded bubble is larger than the contracted one, the mass transfer process is not symmetric. In addition, the diffusive layer is thicker during the contracting phase, resulting in a reduced concentration gradient and slower mass transport than that during the expanded phase. This "shell effect" results in a further net mass flux into the bubble [*Eller and Flynn*, 1965]. There is a net flow of vapor into the bubble, and if bubble growth is limited by the total compressibility of the system, a pressure increase results. Since the pressures in the bubble and the magma are the same except for a small surface tension term, an increase in the pressure inside a bubble translates to an increase in pressure in the entire fluid system. Bubble growth via rectified diffusion in constant pressure systems has been demonstrated by a number of experiments [e.g., *Eller*, 1969] and it is used in industrial applications such as ultrasonic cleaning. However, the increase in pressure in systems where expansion is restricted is a new feature of the formulations of *Sturtevant et al.* [1996] and *Brodsky et al.* [1998]. This modification imposes stringent constraints on the systems in which rectified diffusion can serve as a pressure-building mechanism. If rectified diffusion increases the pressure, then a variety of mechanisms, including failure of the chamber, may ultimately result in an eruption.

The limitations and assumptions of rectified diffusion as a pressure-increasing mechanism in natural magmatic systems are explored here. We present a model of rectified diffusion in magmatic systems which is closely related to the theory of *Sturtevant et al.* [1996] for geothermal systems. We then specify the necessary constraints on the system and numerically estimate the magnitude of the effect. The computed stresses are compared to static stress changes in an elastic halfspace and then we discuss the implications for documented triggering cases.

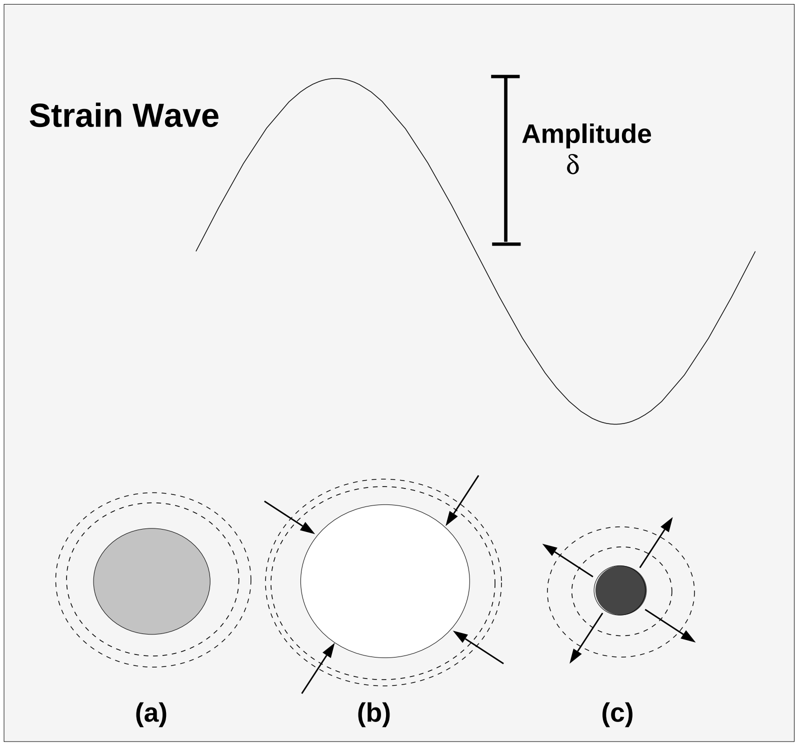

Figure 4.2 Cartoon of rectified diffusion. As the system is excited by a seismic wave, a bubble that is (a) initially in equilibrium is (b) expanded and then (c) compressed. The shading inside the bubble represents the volatile concentration. During the expansion phase (shown in b) the concentration is lower inside the bubble, and hence volatiles diffuse into the bubble as shown by the arrows. During the compression phase (shown in c) volatiles diffuse out of the bubble that has a higher concentration inside. The dashed outer circles represent a shell of constant volume. During expansion the shell is thinner, thus resulting in a faster diffusive flux than when the shell thickens during compression [*Leighton*, 1994]. The diffusive layer surrounding the bubble should be less than the radius of the bubble in order to be in strict accordance with the formulation of *Hsieh and Plesset* [1961]. In this paper the depleted layer is approximately the same size as the bubble.

Model overview

We derive an expression for the increase of pressure produced by a seismic wave passing through a bubbly magma chamber by beginning with an equation of state for the volatiles in the bubble. The strategy of this work is to use this equation of state to write an equation for the evolution of pressure in the bubble (and hence the system) in terms of the evolution of the volume of the bubble and the mass inside the bubble. The mass inside the bubble can then be constrained by the rectified diffusion mass flux of *Hsieh and Plesset* [1961], and the volume of the bubble can be determined by considering the balance in volume changes over the entire magma-volatile system. The result is a rate of pressure increase as a function of the physical parameters of the system.

Theory of rectified diffusion in magmatic systems

For the high pressures and temperatures of a magmatic system, the modified Redlich-Kwong (MRK) equation of state is appropriate for the volatiles [*Holloway*, 1977],

$$\begin{aligned} P &= \frac{RT}{v-b} - \frac{a}{(v^2+bv)T^{1/2}} \\ &= \frac{RT}{V/n-b} - \frac{a}{\left[(V/n)^2+bV/n\right]T^{1/2}}, \end{aligned} \tag{4.1}$$

where P is pressure, T is temperature, V is volume of a bubble, n is number of moles, R is the ideal gas constant, v is the molar volume $(= V/n)$, a is an empirical function of temperature, and b is an empirical constant. Using the MRK equation of state instead of the ideal gas law results in a 20% correction in the specific volumes at the pressures and temperatures of interest. We differentiate (4.1) with respect to time for an isothermal system and arrive at an expression for the rate of pressure change $\dot{P}$ in terms of $\dot{V}$ and $\dot{n}$,

$$\dot{P} = \left(-\frac{\dot{V}}{V} + \frac{\dot{n}}{n}\right) A \tag{4.2}$$

where

$$A \equiv -v \left[\frac{-RT}{(v-b)^2} + \frac{a(2v+b)}{(v^2+bv)^2 T^{1/2}} \right]. \tag{4.3}$$

The system is isothermal over the timescale that diffusion occurs since the diffusion of heat is much more rapid than the diffusion of volatiles. We must evaluate $\dot{n}$ and $\dot{V}$ in order to derive $\dot{P}$ as a function of the state of the system.

We hypothesize that the method for increasing the mass inside the bubbles is rectified diffusion. Note that $\dot{n} = \dot{m}/M_v$ where M_v is the molecular weight of the volatile and $\dot{m}$ is the rate of mass addition by rectified diffusion. *Hsieh and Plesset* [1961] combine the contributions of each half cycle of an oscillating bubble (Figure 4.2) to derive the mass flux of rectified diffusion as

$$\dot{m} = 24\pi D C_s r_0 \delta^2, \tag{4.4}$$

where D is the diffusivity, C_s is the saturated volatile concentration at the mean ambient pressure, r_0 is the bubble radius and δ is the amplitude of the dynamic strain wave. The units of concentration are the same as the units of density (kg/m^3). In the natural system, δ is the strain amplitude of the seismic waves in the magmatic body since most of the compression of the magma-volatile solution occurs in the bubbles. Surface tension is neglected in equation (4.4) and the oscillatory frequency (on the order of Hz) is far from the natural frequency of the bubble (on the order of kHz) [*Brennan*, 1995]. The more rigorous derivation of *Eller and Flynn* [1965] includes surface tension in the definition of C_s. For a solution following Henry's law where the solubility is linearly proportional to pressure, they define

$$C_s' \equiv H_D P \left(1 + \frac{2\sigma}{r_0 P} \right) \tag{4.5}$$

where H_D is the Henry's Law constant and σ is the surface tension. This results in a small, constant correction to the solubility for constant size bubbles. For the magmatic system studied here the correction would be a 0.01% change to H_D when the linear Henry's Law holds. A 0.005% correction is necessary when the solute is

H_2O and the solubility is proportional to the square root of the pressure. In addition, it is possible that surface tension changes during oscillations can cause non-negligible effects in solubility that can affect the total mass flux. The complex details of such a mechanism are the subject of current research by another investigator and beyond the scope of this study [Mie Ichihara, *pers. comm.*].

In order to evaluate the rate of volume change of an individual bubble, $\dot{V}$, one must look at the conservation of volume of the whole system, V_S,

$$\dot{V}_S = \dot{V}_M + N\dot{V}. \tag{4.6}$$

V_M is the volume of the magma, and N is the number of bubbles in the system. The rate of volume change for the entire system $\dot{V}_S$ can be nonzero owing to two distinct processes: (1) leakage of fluid (magma or volatiles) out of the system and (2) deformation of the magma chamber walls. The importance of the first process is calculated by assuming that fluid percolates out of the system into surrounding porous medium following Darcy's law. This volume flux is only significant if a flow to compensate the increased mass of the bubbles can only develop during the passage of the seismic waves. The flow into the porous media develops as fast as the pressure wave diffuses from the wall of the bubble-filled area to the farfield. The pressure wave diffuses from the bubbly area with a diffusion constant κ_p,

$$\kappa_p = \frac{\rho_m k \alpha_m^2}{\eta \phi}. \tag{4.7}$$

where ρ_m is the density of the magma, k is permeability of the surrounding matrix, α_m is the sound velocity of the magma, η is viscosity and ϕ is the porosity [*Philips*, 1991, p. 81]. The pressure reaches approximately the farfield value at a distance L from the wall of the bubble-filled area where L is the radius of the bubble-filled area. Therefore, the porous medium responds to a change in boundary conditions, such as an increase of pressure in the cavity, on the timescale τ_c where

$$\tau_c = \frac{\eta \phi L^2}{\rho_m k \alpha_m^2}. \tag{4.8}$$

As long as τ_c is longer than the duration of shaking Δt, the flow in the porous media cannot respond during the earthquake and a transient pressure rises occurs. The condition that $\Delta t < \tau_c$ is the first major constraint imposed on the magmatic system in order for rectified diffusion to be an effective pressure increasing mechanism. If this condition holds during the earthquake, percolation is negligible and is not relevant to the mass balance in (4.6).

The second process, deformation of the walls, can be evaluated by assuming that the chamber walls are compressed elastically by the increasing pressure in the system. Simplifying the geometry to a sphere,

$$\frac{d_{rr}}{L} = \frac{P}{4G}. \tag{4.9}$$

d_{rr} is the radial displacement and G is the shear modulus of the country rock [*McTigue*, 1987]. The elastic contribution to the change in V_S is

$$\dot{V}_S = \frac{\pi L^3}{G} \dot{P}. \tag{4.10}$$

With increasing pressure, the magma will also compress elastically, and the magma volume change will be

$$\dot{V}_M = \beta V_M \dot{P}, \tag{4.11}$$

where β is the isothermal compressibility of the magma.

Combining (4.2), (4.6), (4.10) and (4.11),

$$\dot{P}\left(1 + \frac{\frac{3}{4}\frac{1}{G} - \beta(1-\phi)}{\phi} A\right) = \frac{\dot{n}}{n} A. \tag{4.12}$$

Equation (4.12) uses the definition of porosity ϕ,

$$\phi = \frac{NV}{V_S}. \tag{4.13}$$

If second term in the brackets of equation (4.12) is much less than 1, the volume change of the bubbles will be negligible. This criterion in terms of porosity ϕ is

$$\phi \gg \phi_{th} \equiv \frac{A(\frac{3}{4}\frac{1}{G} - \beta)}{1 - A\beta}, \tag{4.14}$$

where ϕ_{th} is the threshold porosity. Equation (4.14) explains why a single small bubble does not increase the pressure in a chamber. If there are too few bubbles, the compressibility of the magma and the surrounding rock allows the bubbles to change in size and the resulting pressure change is small. The exact values of the parameters in equation (4.14) depend on the chemistry of the volatiles and the ambient pressure. The elastic moduli $-\beta$ and $1/G$ are approximately 10^{-10} Pa^{-1} and a typical value for the constant A is 2×10^2 MPa. For water bubbles at 1.3×10^2 MPa pressure (lithostatic pressure at a depth of 5 km), ϕ_{th} is 0.026. For CO_2 at the same conditions, ϕ_{th} is 0.037. Temperature and composition dependence in both cases is fairly weak. Since we have already assumed the existence of bubbles, exceeding these small thresholds is not a difficult additional constraint. For the remainder of this work we will assume that the porosity ϕ exceeds the threshold porosity ϕ_{th}.

If the porosity of the system satisfies (4.14), then (4.12) simplifies to

$$\dot{P} = \frac{\dot{n}}{n} A. \tag{4.15}$$

We linearize (4.15) by assuming that departures from the initial state are small, i.e., $A = A_0$ and $n = n_0$. Assuming a spherical bubble, the initial number of moles n_0 is given by

$$n_0 = \frac{4\pi r_0^3 \rho}{3M_v}, \tag{4.16}$$

where ρ is the initial density inside the bubble.

Substituting (4.4) and (4.16) into (4.15),

$$\dot{P} = \frac{18 D C_s \delta^2}{r_0^2 \rho} A_0. \tag{4.17}$$

$\dot{P}$ can be easily calculated from (4.17) for any P, T condition. Note that if a and b were zero, the equation of state would reduce to the ideal gas law with $A_0 = P_0$ and (4.17) would become equation (15) of *Sturtevant et al.* [1996]. The right-hand side of (4.17) is constant, and therefore the final pressure attained depends linearly on the duration of ground motion. The appropriate values of a and b are provided by *Holloway* [1977]. The diffusivity D and the concentration C_s can be estimated from laboratory data as a function of pressure and temperature for a given magma-volatile chemistry (Table 4.1). The final pressure depends on the squares of the bubble radius r_0 and dynamic strain δ. Both parameters must be constrained by observations of natural systems.

Table 4.1. Calibration Data for Solubility and Diffusivity Calculations

Variables	References	Pressure Range, GPa	Temperature Range, °C
Value a and b for CO_2 in MRK	*Holloway* [1977]	0.1-1	100-1000
Value a and b for H_2O in MRK	*Holloway* [1977]	0.01-1 ; 0-0.1	20-1000 ; 0-1300
Solubility CO_2 in basalt	*Dixon* [1992] and *Pan et al.* [1991]	< 0.1 ; 1.0, 1.5	1200 ; 1300-1600
Solubility H_2O in basalt	*Dixon* [1995]	0.018-0.8	1100-1200
Solubility H_2O in rhyolite	*Silver et al.* [1990]	< 0.15	850
Diffusivity CO_2 in basalt	*Blank* [1993]	0.05-0.105[a]	350-1050[b]
Diffusivity H_2O in basalt	*Zhang and Stolper* [1991]	1.0[a]	1300-1500
Diffusivity H_2O in rhyolite	*Zhang et al.* [1991]	0.0001[a]	400-850

[a] No pressure dependence is modeled for diffusivity. [b] These experiments were for granitic composition but pass close to the one basalt point of *Zhang and Stolper* [1991] as noted by *Watson* [1994].

Volatile concentration

Hsieh and Plesset [1961] derived equation (4.4) for the mass flux assuming that the solution was saturated. For solutions that are supersaturated or subsaturated the flux due to ordinary diffusion is superposed on the rectified diffusion flux. In a bubbly system with restricted volume changes, the pressure will rise as long as the total mass flux is positive into the bubble even if the ordinary diffusive mass flux is outward, i.e., the solution is subsaturated. The threshold for a pressure rise can be derived by setting the mass flux due to ordinary diffusion equal to the rectified diffusion mass flux. The steady-state ordinary diffusive mass flow rate in the absence of advection is [*Strasberg*, 1961]

$$4\pi D r_o \left(C_\infty - C_s\right), \tag{4.18}$$

where C_∞ is the volatile concentration in the fluid far from the bubble and once again the surface tension term is absorbed into C_s. Therefore, the total mass flux into the bubble from both ordinary and rectified diffusion is

$$4\pi D r_o \left(C_\infty - C_s\right) + 24\pi D C_s r_0 \delta^2 \tag{4.19}$$

and this expression must be positive for the pressure to grow. Rewriting this statement as a threshold in terms of concentration produces

$$\frac{C_\infty}{C_s} > 1 - 6\delta^2. \tag{4.20}$$

Since δ for the dynamic strains considered here is $\sim 10^{-4}$, the last term is negligible and the fluid must be supersaturated in volatiles in order for the mass flow into the bubble to be positive.

As the pressure increases during the passage of the seismic waves, the solubility of the volatiles C_s also increases. As a result, the relative saturation C_∞/C_s decreases and if it drops below the threshold in (4.20), the mass flow into the bubble will cease. Therefore, there is a relationship between the maximum pressure increase and the initial degree of supersaturation. In order to isolate the effects of the rectified diffusive mass flow, as opposed to the ordinary diffusive mass flow we derive this relationship assuming that the pressure change is caused only by the rectified diffusion (Figure 4.3). The scenario requires that the ordinary diffusive flow is ineffective in raising the pressure for some reason. The situations in which this assumption is applicable are discussed in some detail after the derivation.

For CO_2-magma systems it is appropriate to assume that C_s is linearly proportional to pressure P. We define the initial supersaturation x by

$$C_\infty = (1+x)C_s^0. \tag{4.21}$$

Since the initial saturation concentration C_s^0 is proportional to the initial pressure P_0 and the final C_s is proportional to the final pressure $\Delta P + P_0$, the left-hand side of (4.20) can be written as

$$\frac{C_\infty}{C_s} = \frac{(1+x)P_0}{P_0 + \Delta P}. \tag{4.22}$$

The maximum possible pressure rise ΔP is therefore governed by (4.20):

$$\Delta P < P_0 \left(-1 + \frac{1+x}{1-6\delta^2} \right). \tag{4.23}$$

If $\delta^2 \ll 1$ then (4.23) can be approximated by

$$\Delta P < x P_0. \tag{4.24}$$

Equation (4.24) is appropriate for CO_2 in magmatic systems, but the concentration at saturation C_s of water in magmas at low pressures is not linear in pressure. For water dissolved in silicate melts, C_s is better approximated as proportional to $P^{1/2}$ [*Silver et al.*, 1990]. The maximum pressure increase for a given initial supersaturation is therefore slightly higher for water in magma and bounded by

$$\Delta P < P_0 \left[-1 + \left(\frac{1+x}{1-6\delta^2} \right)^2 \right]. \tag{4.25}$$

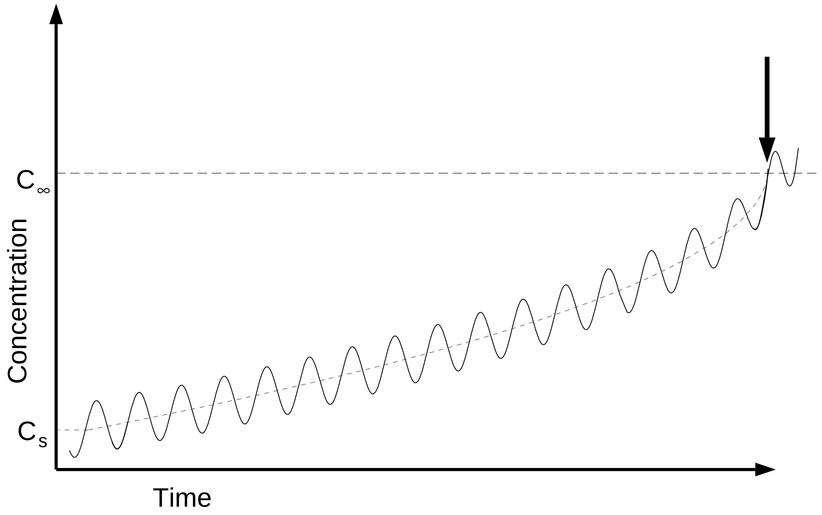

Figure 4.3 Schematic of the change in solubility with pressure due to rectified diffusion alone. The solid sinusoid represents the concentration at the bubble wall. The dashed horizontal line indicates the far-field concentration C_∞, and the dotted curve is the saturation concentration C_s, which increases as the pressure increases. When C_s becomes greater than C_∞ by $\sim 6\delta^2$ at the time indicated by the arrow, the solution becomes subsaturated and rectified diffusion ceases.

Using the same approximation as before, (4.25) to first order in the small parameter x is

$$\Delta P < 2xP_0. \tag{4.26}$$

ΔP is limited in both cases by the initial supersaturation. A 1% supersaturated solution can support approximately a 1% increase in pressure in a CO_2 system or a 2% increase in an H_2O system.

Given the necessity of supersaturation and hence a positive mass flux into the bubble before excitation, one might ask what effect the small increase in mass flux from rectified diffusion would have. Would not the pressure already be rising from diffusive growth in the supersaturated solution? Since the rectified diffusion mass flux is much smaller than the ordinary mass flux, why is it not negligible? To answer these questions in *Brodsky et al.* [1998], we postulated a steady state diffusive system prior to the exciting earthquake. We argued conceptually that in a highly heterogeneous magma body undergoing vigorous convection as first suggested by *Shaw* [1965], there will be some regions that are crystallizing and others that are resorbing mineral phases into the melt at any given time. As the system nears eruption, crystallizing regions that are undergoing second boiling and bubble growth should be common. During the preparatory stages of an eruption there may always be some region of the magma body that has recently become supersaturated and thus has small bubbles present. These regions would be the relevant areas for rectified diffusion to occur. As crystallization slowly supersaturates the magma, a steady-state flow pattern develops and bubbly magma percolates away through the porous medium. When an earthquake occurs, a small, uncompensated increase in mass flux occurs since an increase in the percolative flux cannot develop on this timescale. Even though the rectified diffusion mass flux increase is much smaller than the total mass flux into the bubbles, the rapid excitation prevents the development of a compensating loss of volatiles by percolation. The pressure rises. In *Brodsky et al.* [1998] we never explored the details of the required pre-earthquake bubbly region. Below we propose a specific example of a steady-state diffusive system in order to make the proposed scenario more concrete. In so doing, we are able to quantify the strong constraints that are placed on the areas in which rectified diffusion is applicable.

A pocket or cavity of supersaturated melt surrounded by crystalline magma is shown in Figure 4.4. Here the system has evolved to the steady-state where the volume of the magma percolating out of the cavity balances the volume of the volatiles diffusing into the bubbles. The concentration of volatiles dissolved in the magma inside the bubbly region is maintained in a slightly supersaturated state by continued crystallization. In order to develop the mathematical description of the system we treat the bubbly magma as a continuum fluid with material properites that are a combination of those of the volatile and silicate melt. Conservation of mass at any point inside the cavity requires that

$$\nabla \cdot \mathbf{u} = \frac{D\rho}{Dt} \tag{4.27}$$

where u is the bulk fluid velocity. Neglecting the advective term,

$$\nabla \cdot \mathbf{u} = \frac{\partial \rho}{\partial t}. \tag{4.28}$$

The bulk density is changing due to the degassing of the magma. If the mass fraction of gas is much less than 1,

$$\nabla \cdot \mathbf{u} = \frac{\phi}{V\rho_m} \dot{m}_{OD} \tag{4.29}$$

where $\dot{m}_{OD}$ is the steady-state ordinary diffusive flux,

$$\dot{m}_{OD} = 4\pi D r_0 (C_\infty - C_s). \tag{4.30}$$

For simplicity, the saturation concentration C_s is assumed to be linearly proportional to P over small changes in pressure

$$C_s = H_D P \tag{4.31}$$

where H_D is the Henry's Law constant. By substituting equations (4.30) and (4.31) into equation (4.29), we derive a continuity equation that is coupled to the pressure field,

$$\nabla \cdot \mathbf{u} = \frac{\phi}{V\rho_m} \dot{m}_{OD} = 4\pi D r_0 (C_\infty - H_D P) \frac{\phi}{V\rho_m}. \tag{4.32}$$

The velocity and pressure fields in the magma are also coupled by the Navier-Stokes equation in the creeping flow limit since the Reynolds number is much less than unity,

$$\eta \nabla^2 \mathbf{u} = \nabla P. \tag{4.33}$$

The velocity is assumed to be bounded at $r = 0$. The boundary condition at $r = L$ is $p = p_1$. A solution to the coupled equations (4.32) and (4.33) assuming spherical symmetry is

$$u(r) = u_1 \left(\frac{r}{L}\right)^{1/B} \tag{4.34}$$

$$P(r) = P_1 + \eta \frac{u_1}{L} \frac{1}{B(1-B)} \left[\left(\frac{L}{r} \right)^{1-1/B} - 1 \right] \tag{4.35}$$

where

$$B \equiv \sqrt{1 + \eta H_D 4\pi D r_0 \frac{\phi}{\rho_m V}} \tag{4.36}$$

and

$$u_1 = LB(B-1) \frac{C_\infty - H_D P_1}{H_D \eta}. \tag{4.37}$$

The dimensionless parameter B is approximately 1 for conditions appropriate to a magmatic system. Therefore, the pressure is nearly constant inside the region and the velocity increases almost linearly with r (Figure 4.4). The singularity in pressure at $r = 0$ is unphysical and is asssumed to be smoothed out by the non-continuum processes.

Outside the pressurized cavity is a crystal-rich mush that acts as a porous medium. The bubbly magma exiting the cavity percolates through the mush in accordance with Darcy's law,

$$u = -\frac{k}{\eta} \nabla P \tag{4.38}$$

where k is the permeability. The solution for the flow outside the cavity where $r > L$ is

$$P(r) = P_1 - \frac{u_1 \eta L^2}{k} \left(\frac{1}{L} - \frac{1}{r} \right) \tag{4.39}$$

$$u(r) = u_1 \left(\frac{L}{r} \right)^2 . \tag{4.40}$$

The pressure in the farfield is $P_\infty \equiv P_1 - u_1 \eta L / k$. The cavity pressure has been raised relative to the farfield by an amount equal to $u_1 \eta L / k$. This model makes sense as long as the change in pressure is much less than the farfield pressure, i.e.,

$$\frac{u_1 \eta L}{k} \ll P_\infty. \tag{4.41}$$

In a system that satisfies (4.39)–(4.41), a significant flux exits the cavity to pre-

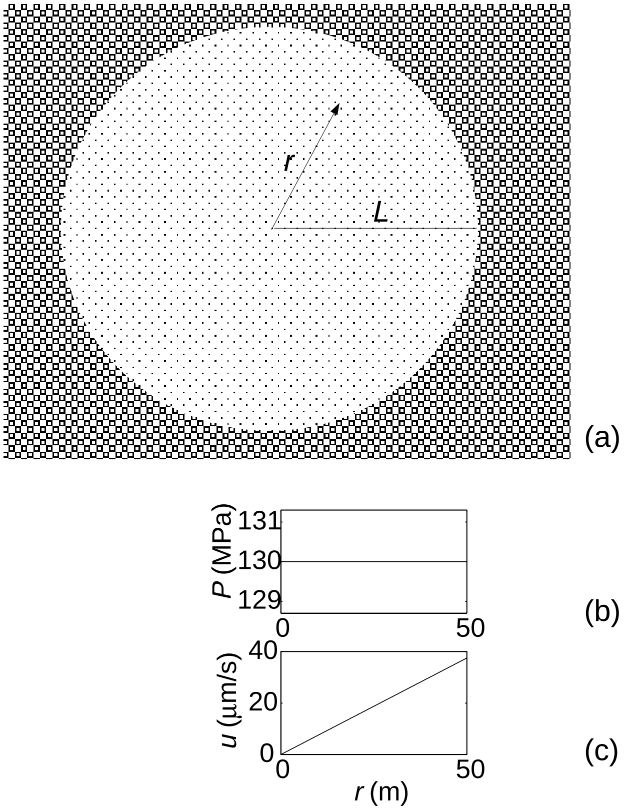

Figure 4.4 (a) Pocket of degassing, bubbly magma (speckled region) with radius L surrounded by crystalline mush (box pattern) with permeability k. Spherical symmetry is assumed with radial coordinate r. Note that r is unrelated to the bubble radius r_0. (b) Pressure in steady state system with parameters in Table 4.2. (c) Velocity in steady state system with parameters in Table 4.2.

vent pressure from building due to rectified diffusion. This steady-state system slowly developed as the magma gradually cooled and crystallization increased the volatile concentration in the magma. The non-equilbrium petrological system must crystallize fast enough to maintain the supersaturation as gas is exsolved into the bubbles. The crystallization cannot proceed so fast as to choke the cavity with crystals during the earthquake. The details of the petrology are beyond the scope of this chapter and may pose additional constraints beyond the ones considered here. As previously discussed, the system responds to a change in boundary conditions, such as an increase of pressure in the cavity, on the timescale τ_c that a wave of pressure can diffuse a distance L from the cavity wall. If τ_c is much greater than the duration of triggering earthquakes, but less than the time over which the ordinary diffusive system developed, then after a large regional earthquake rectified diffusion driven by seismic waves can increase the pressure in the supersaturated system. After substituting (4.37) in to (4.41), the necessary constraints on the parameters of the steady-state system are summarized as:

$$\tau_c > \Delta t \tag{4.42}$$

$$P_\infty > \frac{L^2 B(B-1)(C_\infty - H_D P_1)}{H_D k} \tag{4.43}$$

One set of parameters that satisfies these constraints is in Table 4.2. In this case cavities with 100 m diameter would be pressurized during rectified diffusion. The pressure rise during the earthquake may be much less than the total pressure of the system. However, the pressure increase happens quickly and the rapid change in pressure may be enough to trigger an eruption either directly by stressing the chamber walls or indirectly by increasing the vigor of convection.

This scenario highlights a few of the important requirements for triggering an eruption by dynamic strain. Rectified diffusion is a triggering mechanism. The magma-volatile system must already be present and exsolving gases. The volcanoes studied in this paper would have likely erupted eventually. The regional earthquakes merely accelerated the process. Also, the entire magma body need not be filled with small

Table 4.2. Example Parameter Set

Variables	Values
L	50 m
P_1	1.3×10^8 Pa
η	10^2 Pa s
H_D	10^{-6} kg/m^3/Pa
D	10^{-9} m^2/s
r_0	2×10^{-5} m
ϕ	0.04
ρ_l	2500 kg/m^3
C_∞	$(1 + 10^{-4})\ H_D P_1$
k	10^{-7} m^2
Δt	60 s
α_l	5×10^2 m /s

bubbles in order for rectified diffusion to occur. As long as a region has enough bubbles to meet the porosity constraint in (4.14) and the criteria (4.42)-(4.43) are met, then rectified diffusion will be an effective pressure-raising mechanism. Equations (4.42) and (4.43) are an important extension of *Brodsky et al.* [1998]. The set of parameters used in that paper do not conform to these new criteria. The parameters used in [*Brodsky et al.*, 1998] require $u_1 > 6$ m/s and therefore the pressure in the bubbly region before the earthquake is at least 1000 times the farfield value for any permeability consistent with equation (4.42). Moreover, if u_1 is large, inertial effects will become important and Darcy's law may be inapplicable. Although it is possible that some other steady-state diffusive system exists where the parameters of *Brodsky et al.* [1998] are applicable, the author has been unable to formulate such a system. Therefore, a new set of calculations are presented in this chapter that are consistent with equations (4.42) and (4.43). Equations (4.42) and (4.43) are also applicable to geothermal systems and therefore the work of *Sturtevant et al.* [1996] could be similarly revised.

Another mechanism that was proposed to explain regionally triggered seismicity is subject to similar constraints. *Linde et al.* [1994] suggested that the advective overpressure from rising bubbles can increase the pressure in a magmatic system to unstable levels. As the pressure rises and the solubility increases, the bubbles would tend to resorb. Advective overpressure is also only applicable in a supersaturated system. Unless a steady-state diffusive system exists or the bubbles are stabilized by surfactants [*Leighton*, 1994], the diffusive flux will change the pressure of the system more rapidly than the advective overpressure. At this time no surfactants have been identified in magmatic systems. The search for surfactants may pose an interesting experimental challenge for the future and has dynamical significance for mechanisms such as advective overpressure.

The discussion of the magma-bubble system presented here is more extensive than the previous treatment in *Brodsky et al.* [1998], but it is not exhaustive. Several effects that may further reduce the rectified diffusive mass flow are not treated. For instance, transients in the ordinary diffusive flux during the earthquake due to the solubility change are not included in the co-seismic pressurization model. Such effects would ordinarily be expected to be more significant than rectified diffusion by a factor of $\Delta P/P_0\delta^2$. However, it is possible that they are conteracted by the surface tension changes due to bubble growth. The existence of significant bubble growth on the timescale of the earthquake is implicit in the pre-earthquake ordinary diffusive system. Another effect that would reduce the rectified diffusion mass flux is the depletion of the volatile layer surrounding each bubble. However, the depletion could be negligible since the large fluid velocities in the radial flow modeled in equation (4.40) imply that the volatile-magma solution would be advected over a length greater than the bubble radius on the timescale of the seismic oscillations. The above simplifications are maintained in this treatment not because they are physically justifiable, but rather because the possible reductions in the rectified diffusive flux are not necessary to dismiss rectified diffusion as the best triggering mechanism. As will be shown below, the simplified model here predicts pressure rises comparable or less than static stress changes. It is concluded that rectified diffusion is at best a marginal mechanism. No further calculations are necessary for the purpose of evaluating rectified diffusion as a triggering mechanism.

Physical constraints

The rate of pressure increase is very sensitive to the bubble radius. In a multiple bubble system an effective radius for the entire system can be calculated. The total volume change of all the bubbles present is equal to the volume change of N bubbles of an effective radius r_{eff} [*Sturtevant et al.*, 1996]. In terms of average radius $\overline{r}$ and average cubed radius $\overline{r^3}$, the effective radius is

$$r_{\text{eff}} = \sqrt{\overline{r^3}/\overline{r}}. \tag{4.44}$$

The effective bubble radius is weighted towards the largest bubbles in the distribution. Direct observation of the average bubble size in a magma chamber is impossible. It might be expected that one could learn about bubble sizes from natural volcanic samples. However, the rapid decompression during eruptions dominates the observed bubble size distribution, and no easily discernible information about the *in situ* magma chamber bubble sizes is retained in the rocks [*Cashman and Mangan*, 1994]. *Hurwitz and Navon* [1994] observed average bubble sizes as small as 5 x 10^{-6} m in laboratory nucleation experiments, and *Davis and Ihinger* [1996] observed 10^{-6} m bubbles in similar work. The experimental work provides a lower bound for the size of bubbles a short time after nucleation. We choose to model the effective bubble radius r_{eff} as 20 μm in order to develop a system in compliance with the constraints in equations (4.42)-(4.43). This value is 20 times larger than the one used in [*Brodsky et al.*, 1998]. Since the total pressure accumulated during shaking is proportional to the inverse square of the effective bubble radius, the pressures calculated here are approximately two orders of magnitude lower than those reported in the earlier paper.

The rate of pressure increase is also very sensitive to the amplitude of the dynamic strain wave, δ, at the location of the bubbles. Conventional estimates of the dynamic strain from regional earthquakes are for the measured hard-rock shear wave amplitudes δ_r and therefore are not the relevant δ in the bubbly magma body. Calculating δ from δ_r inevitably involves large uncertainties; we estimate that $\delta = 10$–$15\ \delta_r$ by proposing the following scenario: (1) When shear waves enter a very heterogeneous magma chamber, mode conversion occurs and shear strain is converted to volumetric strain. This process results in some loss of energy but appears to be reasonably efficient since Love waves (transversely polarized shear waves) have been observed with water-well seismographs which are only sensitive to volumetric strain [*Carragan et al.*, 1964]. The scattering also produces high-frequency waves locally. (2) As the waves enter a bubbly magma body, amplification occurs. The amplitude of the wave is increased due to conservation of energy. In addition, a complex series of reflections cause the structure to reverberate. A very simple model of a soft layer between two half-spaces is used to establish a rough overview of how the reverberations would affect the strain amplitude. The Fourier transform of the seismic strain in the magma chamber, $X(\omega)$, is the product of the incident spectrum $I(\omega)$ and the response function $G(\omega)$,

$$X(\omega) = G(\omega)I(\omega). \tag{4.45}$$

In a magma layer of width $2L$ between two hard-rock half-spaces, $G(\omega)$ for a point at a distance x into the soft-layer is

$$G(\omega) = \frac{\alpha_r}{\alpha_m}\frac{1+R}{1-R^2 e^{-4i\omega\frac{L}{\alpha_m}}}\left(e^{-i\omega\frac{x}{\alpha_m}} - Re^{-4i\omega\frac{L}{\alpha_m}}e^{i\omega\frac{x}{\alpha_m}}\right) \tag{4.46}$$

where R is the reflection coefficient

$$R = \frac{\rho_r\alpha_r - \rho_m\alpha_m}{\rho_r\alpha_r + \rho_m\alpha_m}. \tag{4.47}$$

ρ_r and α_r are the hard-rock density and P wave velocity; ρ_m and α_m are the magmatic density and P wave velocity. Laboratory experiments show that melts typically have densities ~5% less than their solid counterparts, and the P wave velocities of melts are ~50% less [*Murase and McBirney*, 1973]. A bubbly magma will have an even lower seismic velocity since the volatile phase is highly compressible [*Kieffer*, 1977].

We use a value of $\alpha_m = 500$ m/s in order to be consistent with the steady-state system in Table 4.2. If $\alpha_r = 6 \times 10^3$ m/s, then the maximum value of the impulse response derived from $G(\omega)$ is 24. Since some energy of the wave is undoubtedly lost in mode conversion and this estimate is approximate, a range of 10–15 is adopted in this work, i.e., the strain amplitude in the magma δ is related to the hard-rock value by

$$\delta = \gamma \delta_r \tag{4.48}$$

where γ is a constant ranging from 10 to 15.

This procedure for determining the amplification factor is admittedly *ad hoc.* Some observational justification for assuming large amplification factors exists. Site amplification of horizontal 0.3–3 Hz velocity waves in the volcanic area of Long Valley results in an increase in amplitude by a factor of 4–5 relative to hard-rock site at Pasadena. Amplification is expected to be even more severe for the high-frequency waves scattered into the slow magmatic system at depth. Surface amplification near small, heterogeneous structures like the San Andreas can be as high as a factor of 10 at 10 Hz. For the 100 m pockets of bubbly magma considered here, the amplification of the high frequency waves is most important. The assumption of large amplification of seismic waves is necessary for rectified diffusion to be effective and is another strong constraint on the systems in which triggering can occur by this mechanism.

The hard-rock value of δ_r for the historical events is estimated by scaling observations of a well-documented earthquake. The hard-rock amplitude of seismic waves is assumed to be the same as that observed at well-placed seismometers. For the regional earthquakes discussed here, we used the strong motion records of the 1985 M=8.1 Michoacan, Mexico, earthquake and the scaling proposed by *Houston and Kanamori* [1990]. This earthquake provides a good analog for the large thrust events that commonly trigger arc volcanoes. The positions of the stations in Michoacan in relation to the fault rupture are similar to the positions of volcanoes along an arc. For very large earthquakes the amplitude of high-frequency waves depends primarily on distance, not magnitude. Since all the earthquakes considered in this work are of sufficient magnitude, it is reasonable to use the scaling

$$\delta_r = \delta_M \left(\frac{\Delta_M}{\Delta} \right)^p, \tag{4.49}$$

where δ_r and δ_M are the strain amplitudes in the earthquake of interest and the Michoacan earthquake, respectively. The distances from the hypocenter to the site being considered are Δ and Δ_M. The exponent p is a scaling factor which is 0.43–0.8 depending on frequency [*Houston and Kanamori*, 1990]. Note that the distance from the volcano to the rupture zone may be considerably less than the distance to the epicenter for large events since M= 8 earthquakes break over 100 km of fault. The scaling of *Houston and Kanamori* [1990] is designed to still be valid in these cases. We used $\Delta_M = 35$ km and $\delta_M = 4.27$ x 10^{-5}, which is the highest amplitude of any wave on this record. Since the pressure increase goes as δ^2, it is only necessary to account for the largest amplitude waves affecting the magmatic system. The dynamic strain δ_r is typically of the order of 10^{-5} for the parameters considered here.

For large earthquakes the duration of the excitation, Δt, scales as the rupture duration $\mathcal{L}/\mathcal{V}$, where $\mathcal{L}$ is the fault length and $\mathcal{V}$ is the rupture velocity. $\mathcal{L}$ is approximately 1000 km for the $M_w = 9.5$ 1960 Chilean earthquake. The largest amplitude waves are observed for about one–half the rupture duration on the Michoacan records. Using the conventional seismic scaling relation $M_w \propto 2 \log \mathcal{L}$, we approximate the duration in seconds,

$$\Delta t = \frac{1}{2} \left(10^{\frac{M_w - 9.5}{2}} \frac{1000}{2.9} \right), \tag{4.50}$$

where the rupture velocity is assumed to be 2.9 km s^{-1}.

Model systems

Equation (4.17) was evaluated for model systems defined by a volatile (carbon dioxide or water) in an end-member magma (basalt or rhyolite) in order to establish the general trends and sensitivities of the rectified diffusion pressure increase. The effects of temperature and pressure are incorporated into the solubility and diffusivity esti-

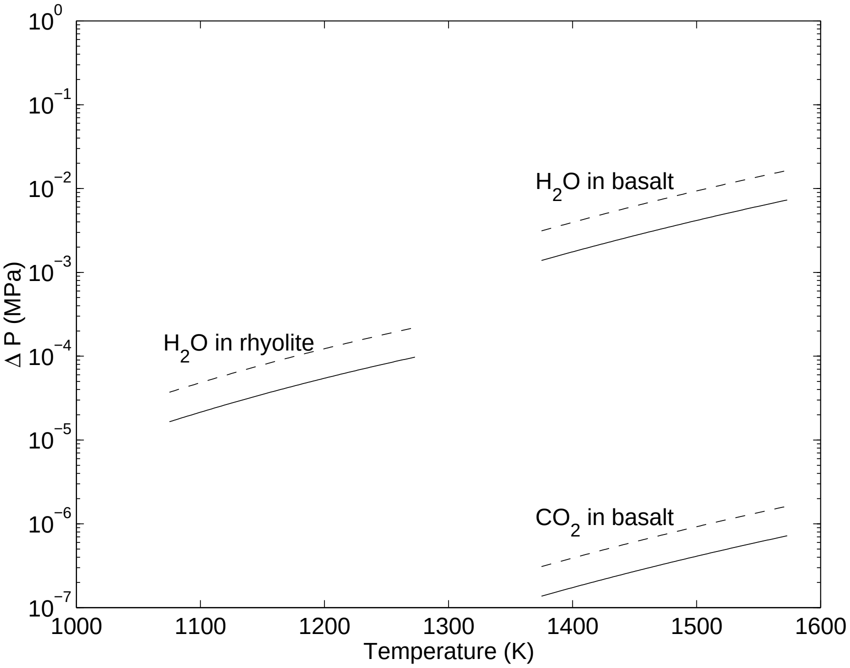

Figure 4.5 Temperature dependence of pressure increase ΔP in model systems via rectified diffusion. The dynamic strain measures 10^{-5} at the surface and lasts 45 s. The magma chamber is at 1.3 x 10^2 MPa (lithostatic pressure at 5 km depth). For each model system, two curves are plotted to show a range of amplification factors (δ/δ_r) of 10-15.

mates from laboratory data (Table 4.1). A range of values are considered, and the sensitivity of the results is shown in Figure 4.5 and Figure 4.6.

The temperature dependence of the pressure increase from rectified diffusion is shown in Figure 4.5; pressure dependence is shown in Figure 4.6. Figures 4.5 and 4.6 demonstrate the relative efficacy of different magma-volatile systems. Since the diffusivity of volatiles in basalt is much higher than that in rhyolite at natural temperatures, rectified diffusion is more effective in basaltic systems. For a basaltic system with water bubbles initially at 130 MPa pressure, the excess pressure from rectified diffusion is between 0.001 and 0.02 MPa. Figures 4.5 and 4.6 show that when CO_2 is the dominant volatile in basaltic systems, the effect of rectified diffusion is very small. This difference arises because CO_2 is much less soluble in magmas than H_2O; that is, C_s is lower. Since the diffusivity of CO_2 in rhyolite is even lower than that of CO_2 in

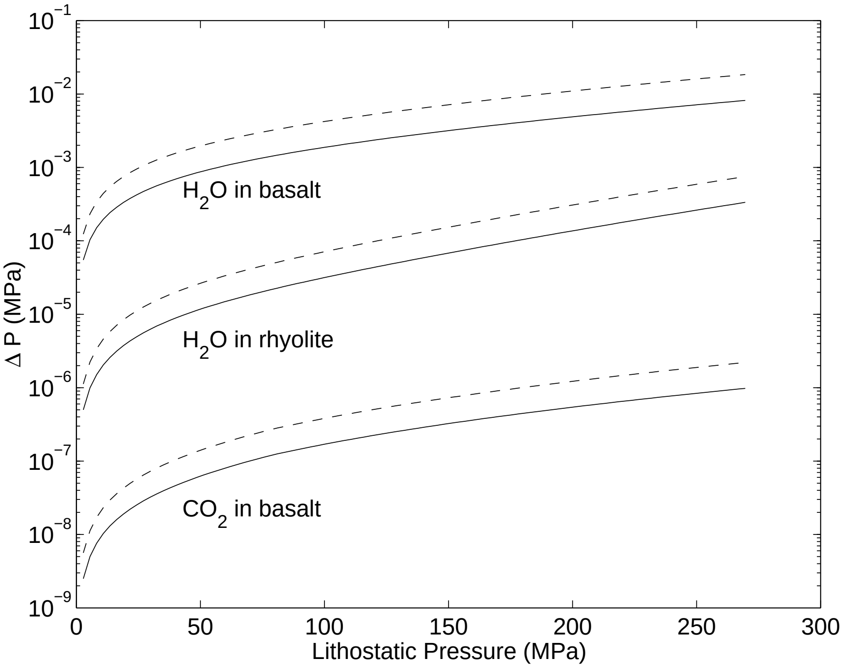

Figure 4.6 Pressure dependence of pressure increase ΔP in model systems via rectified diffusion. The basalt is modeled at 1450 K and the rhyolite is modeled at 1200 K. The excitation and ranges of amplification are the same as those in Figure 4.5.

basalt [*Watson*, 1994] and the solubility is ~30% lower in rhyolite [*Blank*, 1993], the pressure increase will be even smaller than in the CO_2-basalt system. Therefore, the CO_2-rhyolite system is not considered further in this work. In all cases the pressure rise is less than 0.01% of the ambient pressure and so would be possible in a slightly saturated system like the one described in Table 4.2.

Historical cases

Once a pressure increase occurs, a number of processes can eventually lead to dike propagation and a volcanic eruption. The nature of these mechanisms is uncertain and therefore the appropriate threshold for triggering an eruption is currently unknown. The goal of the present work is to compare the pressure increase from rectified diffusion to the stresses produced by the most widely discussed mechanism, static stress changes. After the magma body is pressurized, the eruption can occur at any time. The pressure on the bubbly region walls may cause failure and induce an eruption immediately. Alternatively, the increased pressure gradients outside the bubbly region may sufficiently destabilize the magmatic system to begin a series of events that lead to an eruption some indefinite amount of time later.

A number of instances of large tectonic events triggering volcanic eruptions have been documented. Demonstrating triggering is at best a subjective process. Occasionally, there is a particularly spectacular example such as the 1835 $M = 8.5$ Chile earthquake. Immediately after the earthquake, four eruptions began simultaneously. More often, the relationship is less clear. For this study, only eruptions occurring within 10 days of a large regional earthquake greater than 100 km from the volcano are considered. Table 4.3 lists some of the best-documented cases, and Table 4.4 shows the relevant modeling parameters. Locations are mapped in Figure 4.1.

Table 4.3. Documented Triggered Eruptions

Event Number	Volcano	Year	Distance, km	Delay, days	VEI	Earthquake Magnitude	SGVN
1	Robinson Crusoe[a]	1835	635	0	1?	8.5	1506-02=
2	Minchinmavida[a]	1835	664	0	2	8.5	1508-04=
3	Cerro Yanteles[a]	1835	733	0	2	8.5	1508-051
4	Peteroa[a]	1835	283	?	2	8.5	1507-04=
5	Liamuiga[b]	1843	116	0	?	8.2	1600-03=
6	Unnamed[b](15.97deg N, 61.43deg W)	1843	101	9	?	8.2	1600-07=
7	Llullaillaco[c]	1877	510	?	2	8.5	1505-11=
8	Ambrym[d]	1950	237	2	4	8.1	0507-04=
9	Karpinsky group[d]	1952	404	1	1	8.3	0900-35=
10	Tao-Rusyr Caldera[d]	1952	501	8	3	8.3	0900-31=
11	Puyahe[e]	1960	226	2	3	9.5 M_w	1507-141

Earthquake magnitudes are M_s except where otherwise indicated and are estimated from intensities for preinstrumental events. The event numbers correspond to Figures 4.1 and 4.7. Distances are from epicenter to volcano. The delay is the time between the earthquake and the eruptive event. The volcanic explosivity index (VEI) and catalog number from the Smithsonian Global Volcanism Network (SGVN) are included for reference. References: [a]*Darwin* [1896]. [b]*Robson and Towblin* [1966]. [c]*Casertano* [1963]. [d]*Newhall and Dzurisin* [1988]. [e]*Barrientos* [1994].

Table 4.4. Calculated Values For Triggered Eruptions

Event Number	Volcano	δ_1 $\times 10^5$	δ_2 $\times 10^5$	Δt, s	Petrology	Model	ΔP, MPa
1	Robinson Crusoe	1.2	0.42	55	basalt^{a,b}	b	7.2×10^{-4}–1.3×10^{-2}
2	Minchinmavida	1.2	0.41	55	basalt^{a,c}	b	6.9×10^{-4}–1.3×10^{-2}
3	Cerro Yanteles	1.2	0.37	55	basalt^{a,d}	b	5.6×10^{-4}–1.3×10^{-2}
4	Peteroa	1.7	0.80	55	mafic andesite^{a,e}	b	2.6×10^{-3}–2.7×10^{-2}
5	Liamuiga	2.6	1.6	49	basalt and andesite^{f}, phreatic	b	9.3×10^{-3}–5.5×10^{-2}
6	Unnamed	2.7	1.8	49	geyser ?		
7	Llullaillaco	1.3	0.50	55	dacite^{g}	r	1.3×10^{-5}–2.0×10^{-4}
8	Ambrym	1.9	0.92	35	basalt^{h}	b	2.2×10^{-3}–2.1×10^{-2}
9	Karpinsky group	1.5	0.60	44	andesite^{i}		
10	Tao-Rusyr Caldera	1.4	0.5	44	andesite^{j}		
11	Puyahe (Cordon Calle)	1.9	0.96	173	rhyolite/rhyodacite^{k}	r	1.5×10^{-4}–1.3×10^{-3}

The strains δ_1 and δ_2 are upper and lower limits on dynamic strain outside the magma body from equation (4.49), Δt is the duration of the strong shaking, petrology is the type of observed eruptive products, model indicates the simplified system used (either basalt,"b," or rhyolite, "r") and ΔP is the pressure increase due to rectified diffusion. Notes: [a]The products were extrapolated from the general trends for the volcano. [b]*Casertano* [1962]. [c]*Casertano* [1963]. [d]*Fuenzalida and Etchart* [1974]. [e]*Tormey et al.* [1989]. [f]*Robson and Towblin* [1966]. [g]*DeSilva and Francis* [1991]. [h]*McCall et al.* [1970]. [i]*Mooser et al.* [1958]. [j]*Newhall and Dzurisin* [1988]. [k]*Gerlach et al.* [1988].

Most of the earthquake-volcano systems can be approximated by one of the model cases shown in Figures 4.5 and 4.6. Where such an approximation was not possible, such as for the andesitic systems where no laboratory data are available, no further calculation was attempted. The calculated increase in pressure for all cases is less than 0.04%. Therefore, the systems require at most 0.02% supersaturation in the bubbly region of the magma body in order to maintain rectified diffusion throughout the duration of shaking Δt in accordance with (4.26).

The static stress changes $\Delta\sigma_s$ is computed for comparison by using a double-couple point source model for an earthquake,

$$\Delta\sigma_s = \frac{M_0}{4\pi r^3}, \tag{4.51}$$

where r is the distance from the volcano to the hypocenter and M_0 is the seismic moment. Equation (4.51) is the stress in the direction of maximum amplitude and so is an upper bound on static stress change. Therefore, the comparison with rectified diffusion presented in Figure 4.7 is conservative. The pressure increases from rectified diffusion presented here are approximately two orders of magnitude less than those reported in *Brodsky et al.* [1998]. The difference is a result of the larger bubble radius that was used in order to be consistent with the new constraints in equations (4.42) and (4.43).

Figure 4.7 shows that in most triggering cases the pressure increase by rectified diffusion ΔP is nearly equal to the static stress change $\Delta\sigma_s$. In the silica system ΔP is negligible. The most noticeable difference between rectified diffusion and static stress is that ΔP is nearly independent of the epicentral distance, whereas $\Delta\sigma_s$ decreases quickly with increasing distance. This is because the surface waves that are driving rectified diffusion decay as approximately $r^{-1/2}$ whereas the static stress decays as r^{-3}. For the eruption of Puyahe following the extremely large 1960 $M_w = 9.5$ Chile earthquake, the volcano is effectively in the near field and it is likely that the direct deformation effects are important [*Barrientos*, 1994]. On the other had, rectified diffusion provides at least as good an explanation as static stress changes for the

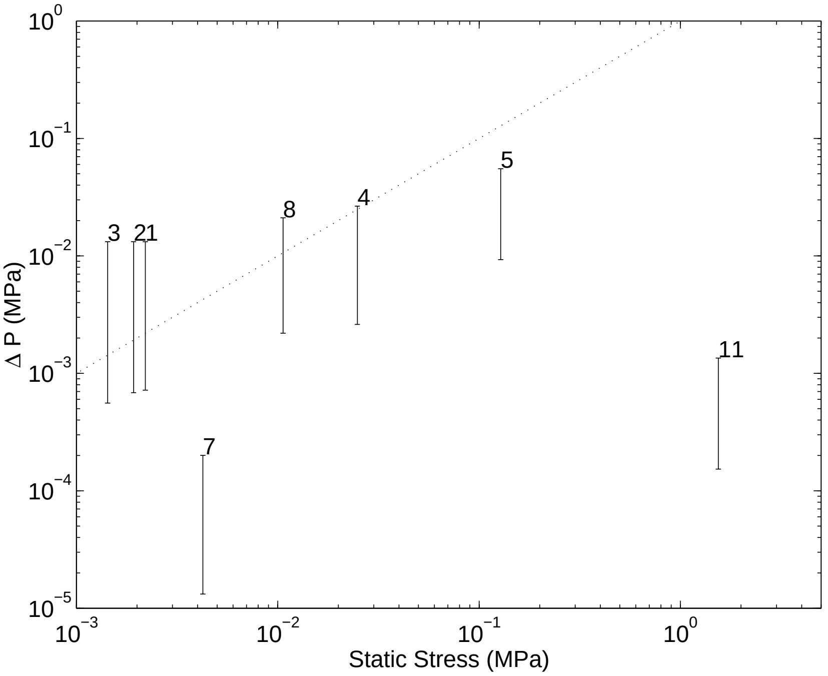

Figure 4.7 A comparison of the stresses from static strain, $\Delta\sigma_s$, and rectified diffusion, ΔP, for cases where a simple model system was applicable. Elastic stress change is calculated from equation (4.51) and plotted on the x axis. The minimum and maximum rectified diffusion pressures (ΔP) are calculated for each case at 1.3 x 10^2 MPa with a temperature of 1473 K for basalt and 1173 K for rhyolite. The results are plotted on the y axis. The range from maximum to minimum takes into account a range of δ corresponding to a range in possible hard-rock strains as estimated in equation (4.49) and amplification factors. Since the results are most sensitive to the estimates of δ and r_0, these are the largest sources of error in the calculations, and therefore the range shown can be taken as a crude estimate of the error bars. Results above the dashed line indicate that ΔP is larger than the static stress from the elastic model. Numbers correspond to the event numbers in Table 4.3.

more distant triggering events such as Cerro Yanteles.

Conclusions

The above calculations indicate that rectified diffusion is a marginally viable mechanism for explaining distant seismic triggering of volcanic eruptions. Strong constraints have been placed on the magma-bubble system. Small (∼100 m) regions of bubble-filled magma must be slightly (∼0.01%) supersaturated before the earthquake. The total porosity must be at least a few percent, and the radii of individual bubbles are assumed to be on the order of 10s of microns. Flow out of these regions into the surrounding partially molten system must establish a delicately balanced steady state where the volume flux out of the cavity exactly balances the volatile flux exsolving from the supersaturated magma. The supersaturation must be maintained by rapid crystallization. The diffusive time for a wave of fluid pressure in the permeable crystalline mush must be long enough that the flow cannot readjust during the passage of the seismic waves, i.e., $\tau_c > 60$ s. The velocity structure must be highly heterogeneous such that it scatters the seismic waves, promotes mode conversion, and amplifies compressional waves by about 10–15 times.

Given a system that is consistent with these strong constraints, rectified diffusion may be as effective a triggering mechanism as static stress in mafic systems. It can be a slightly better explanation for the most distant cases of triggered eruptions. It must be remembered that the calculation of ΔP is based on a simplified model system. Effects such as rupture directivity and multicomponent volatile species have not been included. The inherent uncertainties in estimating physical parameters like bubble size and the dynamic strain amplitude prevent a more sophisticated treatment of such processes from being quantitatively useful.

The stresses produced by both rectified diffusion and static deformation are extremely small. The magmatic system must be very delicately balanced if a 10^{-2} MPa change in stress triggers an eruption. Neither mechanism is completely satisfactory in this regard and this evaluation of rectified diffusion suggests that the search for an appropriate regional triggering mechanism is not finished. However, we have been able to make important progress in outlining some of the basic physics of how seismic waves interact with complex, multiphase magmatic system. It is likely that any successful explanation of the observations will require the incorporation of such fluid dynamical mechanisms into seismological studies. Here we have performed a case study of one mechanism and have constructed a template for future studies of the interactions of seismic waves with magmatic fluids.

References

Barrientos, S., Large thrust earthquakes and volcanic-eruptions, *Pure Appl. Geophys.*, *142*, 225–237, 1994.

Blank, J., An experimental investigation of the behavior of carbon dioxide in rhyolitic melt, Ph.D. thesis, Calif. Institute. of Technol., 1993.

Brennan, C., *Cavitation and Bubble Dynamics*, Oxford Univ., Oxford, UK, 1995.

Brodsky, E. E., B. Sturtevant, and H. Kanamori, Earthquake, volcanoes and rectified diffusion, *J. Geophys. Res.*, *103*, 23,827–23,838, 1998.

Carragan, E., F. Michalko, and S. Katz., Water wells in earthquake and explosion detection, *Tech. rep.*, Off. of Aerosp. Res., U.S. Air Force, Bedford, Mass., 1964.

Casertano, L., *Catalogue of the Active Volcanoes of the World, Including Solfatara Fields: Part XV Chilean Continent*, International Volcanological Association, Napoli, Italy, 1962.

Casertano, L., General characteristics of active andean volcanoes and a summary of their activities during recent centuries, *Bull. Seism. Soc. Amer.*, *53*, 1415–1433, 1963.

Cashman, K., and M. Mangan, Physical aspects of magmatic degassing. 2. constraints on vesiculation processes from textural studies of eruptive products, *Volatiles in Magmas*, *30*, 447–478, 1994.

Darwin, C., *Journal of Researches Into the Natural History and Geology of Countries Visited During the Voyage of H.M.S. Beagle Round the World*, Appleton, New York, 1896.

Davis, M. J., and P. D. Ihinger, Crystal nucleation on bubbles in hydrous silicate melt, *Eos*, *77*, F819, 1996.

DeSilva, S., and P. Francis, *Volcanoes of the Central Andes*, Springer-Verlag, New York, 1991.

Dixon, J., Water and carbon dioxide in basaltic magmas, Ph.D. thesis, Calif. Inst. of Tech., 1992.

Dixon, J., An experimental study of water and carbon dioxide solubility in mid-ocean ridge basaltic liquids, part I: Calibration and solubility models, *J. Petrol.*, *36*, 1607–1631, 1995.

Eller, A., Growth of bubbles by rectified diffusion, *Jour. Acoust. Soc. Am.*, *46*, 1246–1250, 1969.

Eller, A., and H. Flynn, Rectified diffusion during nonlinear pulsations of cavitation bubbles, *J. Acoust. Soc. Am.*, *37*, 493–503, 1965.

Fuenzalida, P., and H. Etchart, Evidencias de migracion volcanica recient desde la linea de volcanes de la patagonia chilena, in *Proc. Sympos. Andean and Antarctic Volcanology Problems*, p. 392, 1974.

Gerlach, D., F. Frey, H. Moreno-Roa, and L. Lopez-Escobar, Recent volcanism in the Puyahue-Cordon Caulle Region, Southern Andes, Chile (40.5°S): Petrogenesis of evolved lavas, *J. of Petrology*, *29*, 333–382, 1988.

Holloway, J., *Thermodynamics in Geology*, chap. Fugacity and activity of molecular species in supercritical fluids, pp. 161–180, D. Reidel, Norwell, Mass., 1977.

Houston, H., and H. Kanamori, Comparison of strong-motion spectra with teleseismic spectra for three magnitude 8 subduction-zone earthquakes, *Bull. Seism. Soc. Amer.*, *81*, 913–934, 1990.

Hsieh, D.-Y., and M. S. Plesset, Theory of rectified diffusion of mass into gas bubbles, *J. Acoust. Soc. Am.*, *33*, 206–215, 1961.

Hurwitz, S., and O. Navon, Bubble nucleation in rhyolitic melts: Experiments at high pressure, temperature, and water contents, *Earth Plan. Sci. Let.*, *122*, 267–280, 1994.

Johnson, M., A. Anderson, and M. Rutherford, Pre-eruptive volatile contents of magmas, *Volatiles in Magmas*, *30*, 281–330, 1994.

Kieffer, S., Sound speed of liquid-gas mixtures: Water-air and water-steam, *J. Geophys. Res.*, *82*, 2895–2904, 1977.

Kimura, M., *Geodynamics of the Western Pacific*, chap. Significant eruptive activities related to large interplate earthquakes in the Northwestern Pacific margin, pp. 557–570, Jpn. Sci. Soc., Tokyo, 1978.

Lambert, G., M. Cloarec, B. Ardouin, and J. LeRoully, Volcanic emission of radionuclides and magma dynamics, *Earth Plan. Sci. Let.*, *76*, 185–192, 1985.

Leighton, T., *The Acoustic Bubble*, Academic, San Diego, Calif., 1994.

Linde, A., and I. Sacks, Triggering of volcanic eruptions, *Nature*, *395*, 888–890, 1998.

Linde, A., I. Sacks, M. Johnston, D. Hill, and R. Bilham, Increased pressure from rising bubbles as a mechanism for remotely triggered seismicity, *Nature*, *371*(6496), 408–410, 1994.

Lowenstern, J. B., *Magmas, Fluids, and Ore Deposits*, chap. Applications of silicate-melt inclusions to the study of magmatic volatiles, pp. 71–100, Miner. Assoc. of Can., 1995.

McCall, G., R. LeMaitre, A. Malahoff, G. Robinson, and T. Stephenson, The geology and geophysics of the Ambrym Caldera, New Hebrides, *Bull. Volcanol.*, *34*, 681–696, 1970.

McTigue, D., Elastic stress and deformation near a finite spherical magma body: Resolution of the point source paradox, *J. Geophys. Res.*, *92*, 12,931–12,940, 1987.

Mooser, F., H. Meyer-Abich, and A. McBirney, *Catalogue of the Active Volcanoes of the World, Including Solfatara Fields: Part V Melanesia*, International Volcanological Association, Napoli, Italy, 1958.

Murase, T., and A. McBirney, Properties of some common igneous rocks and their melts at high temperatures, *Geol. Soc. Am. Bull.*, *84*, 3563–3592, 1973.

Newhall, C., and D. Dzurisin, Historical unrest at large calderas of the world, *U.S. Geol. Surv. Bull.*, *1855*, 1988.

Pan, V., J. Holloway, and R. Hervig, The pressure and temperature dependence of carbon dioxide solubility in tholeiitic basalt melts, *Geochim. Cosmochim. Acta*, *55*, 1587–1595, 1991.

Philips, O. M., *Flow and Reactions in Permeable Rocks*, Cambridge Univ., Cambridge, 1991.

Robson, G., and J. Towblin, *Catalogue of the Active Volcanoes of the World, Including Solfatara Fields: Part XX West Indies*, International Volcanological Association, Napoli, Italy, 1966.

Shaw, H., Comments on viscosity, crystal settling, and convection in granitic magmas, *Am. J. Sci.*, *263*, 120–152, 1965.

Silver, L., P. Ihinger, and E. Stolper, The influence of bulk composition on the speciation of water in silicate glasses, *Contrib. Mineral. Petrol.*, *104*, 142–162, 1990.

Strasberg, M., Rectified diffusion: comments on a paper of Hsieh and Plesset, *J. Acoust. Soc. Am.*, *33*, 259, 1961.

Sturtevant, B., H. Kanamori, and E. E. Brodsky, Seismic triggering by rectified diffusion in geothermal systems, *J. Geophys. Res.*, *101*, 25,269–25,282, 1996.

Tormey, D., F. Frey, and F. Escobar, Geologic history of the active Azufre-Planchon-Peteroa volcanic center (35°15’s, southern Andes), with implications for the development of compositional gaps, *Assoc. Geol. Argent. Rev.*, *44*, 420–430, 1989.

Watson, E. N., Diffusion in volatile-bearing magmas, *Volatiles in Magmas*, *30*, 371–411, 1994.

Yamashina, K., and K. Nakamura, Correlations between tectonic earthquakes and volcanic activity of Izu-Oshima volcano, Japan, *J. Volcanol. Geotherm. Res.*, *4*, 233–250, 1978.

Zhang, Y., and E. Stolper, Water diffusion in a basaltic melt, *Nature*, *351*, 306–309, 1991.

Zhang, Y., E. Stolper, and G. Wasserburg, Diffusion of water in rhyolitic glasses, *Geochim. Cosmochim. Acta*, *55*, 441–456, 1991.

Part III

Seismic Determination of Mass Ejection Rates

Chapter 5 A Seismically Constrained Mass Discharge Rate for the Initiation of the May 18, 1980, Mount St. Helens Eruption

Abstract

We calculate the vertical mass discharge rate from Mount St. Helens for the first few minutes of the May 18, 1980 cataclysmic eruption using a new method based on seismic constraints. The observed seismic waves indicate that the seismic source is a series of single forces. We model these forces as thrusts due to a combination of the momentum flux of the erupted products and the pressure of the eruptive jet. The momentum discharge rate is converted to a mass discharge rate based on estimates of the velocity and jet pressure as constrained by a simple fluid dynamical model. Only two parameters are necessary for the calculation: Mach number and sound velocity. The calculated mass ejected in the first 100 s is 1.6×10^{11}–4.6×10^{11} kg. Since the total blast deposit is $\sim3.2\times10^{11}$–4.1×10^{11} kg, one possible interpretation is that the directed blast had a significant ($\geq$40%) vertical component.

Introduction

One of the most fundamental measures of explosive volcanic eruptions is mass discharge rate. Quantification of the rate and orientation at which mass is ejected improves the observational constraints on the eruptive process and provides realistic hazard assessment. In this paper, we combine a fluid mechanical model with seismological data to formulate a method for measuring mass emission rate.

The most common previously established method of assessing the volcanic mass emission rate is measurement of the column height. *Wilson et al.* [1978] show that the height of a plinian column is a function of the energy available for buoyant ascent. Since the energy source is hot mass ejected from the vent, height can be inverted for mass emission rate. This method has proven effective for computing average discharges of plumes [*Sparks et al.*, 1997] but is not applicable for certain eruptive processes including directed blasts, strombolian explosions, and pyroclastic flow generation. In all of these cases the erupted mass of interest does not directly contribute to the buoyant plume and therefore must be measured by other means. Another shortcoming of the column height method is that time resolution is limited since column height is usually reported as an average value over several hours. Our method provides both the time history of mass emission rate and the orientation of the jetting using continuously measured seismic data. Such resolution allows us to quantify the sequence of events initiating an eruption.

We illustrate the new method by studying the cataclysmic eruption of Mount St. Helens. On the morning of May 18, 1980, a giant landslide exposed the cryptodome beneath the north flank of Mount St. Helens. The hot, pressurized magma exploded and over the next few minutes 3–4×10^{11} kg of material surged northward over an area of approximately 600 km^2. Although asymmetric explosive eruptions had been previously identified and termed "directed blasts" [*LaCroix*, 1930; *Gorshkov*, 1963], the devastation at Mount St. Helens brought the eruptive style to the attention of the volcanological community.

Many studies using a variety of data sets have addressed the Mount St. Helens blast, but none were able to include direct measurement of the mass emission rates. For the following practical and scientific reasons such a measurement would be useful:

1. Measurements of erupted mass quantify the hazard. Both the rate of eruption and the total mass of the products are useful measures of the size of an eruption as noted by *Walker* [1980]. Total erupted mass is a major criterion for the Volcanic Explosivity Index (VEI) [*Newhall and Self*, 1982] and such metrics constitute an important tool for formulating effective hazard management policies.
2. Various eruptive styles can be distinguished by their mass discharge rates. At Mount St. Helens the blast was unexpectedly and disproportionately devastating. This phenomenon was presumably due to the important role of momentum, or in the terminology of *Walker* [1980], the "violence" of this particular eruptive style. Measuring and documenting the mass emission rate of eruptions can clarify such features of eruptive styles.
3. Measurement of the erupted mass relates eruptive events to their products. The total mass in a unit can be estimated by mapping the deposits. Providing a history of the mass ejection over the course of the eruption could potentially associate the deposits with the processes that produced them. Seismic data can constrain the geometry of the source in addition to its time history.
4. Mass ejection rate $\dot{M}$ provides a direct constraint on dynamic models of an eruption. A number of fluid mechanical models of explosive eruptions have been developed [e.g., *Wilson et al.*, 1980; *Sparks et al.*, 1997] and the resulting flow solutions include predictions of mass flux. Accurate determination of $\dot{M}$ can test and calibrate these models.

The inversion for $\dot{M}$ begins with a series of equivalent forces that were previously found from the seismic data for May 18, 1980 [*Kanamori and Given*, 1982; *Kanamori et al.*, 1984]. Seismic data have been analyzed to produce similar equivalent force systems during eruptions at other volcanoes [*Uhira et al.*, 1994; *Uhira and Takeo*, 1994; *Nishimura*, 1995]. In this paper, we model the observed seismic pulses as thrusts due to the momentum flux of the erupted products. The momentum discharge rate is then converted to a mass discharge rate based on an independent estimate of the velocity. The resultant mass discharge rate is integrated over time to calculate the total mass corresponding to the observed seismic forces. We then reinterpret the eruption in light of this new data set. Our calculations allow us to evaluate the directionality of the blast.

The model used to connect the seismic observations with dynamical quantities is deliberately simplified to provide an analytical method of evaluating the data. As such, it inevitably omits features of a complex volcanic flow such as multiphase flow effects and fragmentation dynamics. Wherever possible, we justify the omissions through quantitative assessment. Such assessments are necessarily approximate and result in generous error estimates. Numerical treatment of the full solid and fluid dynamical systems could refine the results presented here and would be a logical extension of this work.

Observations

At 1532:11 UT (0832:11 LT) on May 18, 1980, seismic activity increased at Mount St. Helens and the cataclysmic eruption began. Within the next few minutes three separate landslide blocks detached and dark plumes were photographed rising from the summit. About 40 seconds after the activity began, the largest slide block had reached the bottom of the mountain and a sudden expansion of light-colored material occurred. A light-colored, ground-hugging flow is first visible in photographs at this time [*Hickson*, 1990]. These rolling clouds are what is commonly referred to as the directed blast and have been variously interpreted as an underexpanded jet, pyroclastic flow or surge [*Kieffer*, 1981; *Walker and McBroome*, 1983; *Waitt*, 1984; *Hoblitt and Miller*, 1984]. The resulting deposit is called unit A and is subdivided into sedimentary layers A0–A3 [*Waitt and Dzurisin*, 1981; *Hoblitt*, 1989]. A giant mushroom cloud rose from the devastated area about 4 min after the seismic activity increased [*Sparks et al.*, 1986]. The rapid ascent of the cloud was measured in satellite photos [*Sparks et al.*, 1986] and reported by airline pilots [*Rosenbaum and Waitt*, 1981]. The fallout of this mushroom cloud formed layer A3, a fine, dark deposit distributed over the entire devastated area. A3 contains abundant shredded vegetation which suggests that the material flowed along the ground before being lofted into the mushroom cloud and eventually deposited [*Sisson*, 1995]. The next layer deposited (B1) was a “salt and pepper” combination of lithics and pumice [*Sarna-Wojicki et al.*, 1981]. B1 marks the introduction of juvenile pumice into the deposits. It lies above A3 almost everywhere and the boundary between the units is sharp [*Waitt and Dzurisin*, 1981]. This indicates that some time elapsed between the mushroom cloud and any pumiceous Plinian air fall. There was no Plinian column from the vent associated with the blast phase of the eruption [*Criswell*, 1987]. This important observation is reinforced by satellite images that show the blast column from the vent only reached a height of $\sim$6 km above the mountain rather than expanding buoyantly into the upper atmosphere [*Sparks et al.*, 1986]. Therefore, the dynamics of the blast initiation can best be understood by studying the compressible, momentum driven flow in the gas-thrust region [*Sparks et al.*, 1997] rather than by pursuing thermally driven plume theory.

These eruptive events were also observable seismically. Explosions and flows coupled to the ground to generate seismic waves. The various phases on the seismograms suggest distinct processes that can be modeled by systems of equivalent forces. Each process is distinguished by the frequency content and the geometry of the source (Figure 5.1a). Figure 5.2 reviews the evidence for a series of vertical single forces as first documented by *Kanamori et al.* [1984]. The near-field station at Longmire (Figure 5.2a) recorded the ground motion of the source directly and simple visual inspection suggests a source of the form illustrated in Figure 5.1b. The teleseismic data (Figure 5.2b) both characterizes the type of equivalent force system of the seismic source and quantifies the orientation and amplitude of those forces. The key observation for determining the force system is the azimuthal invariance of the P waves (Figure 5.2b). The nearly identical amplitudes and waveforms at all azimuths strongly suggest vertical single force sources. This pattern of radiated energy is incompatible with the four-lobed pattern of the double couple force model used for most tectonic earthquakes. Furthermore, the large amplitudes of the P waves relative to

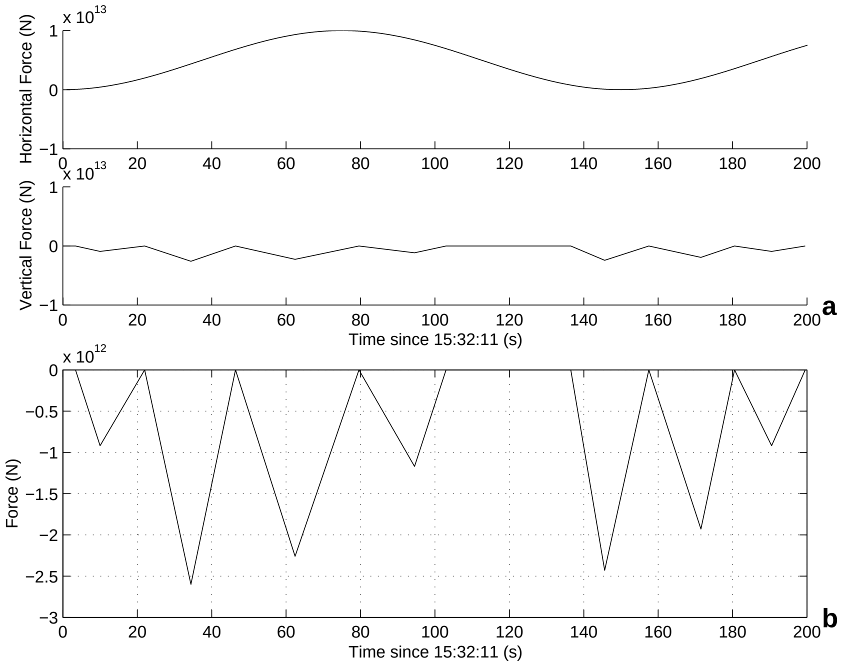

Figure 5.1 Time history of the seismic sources. (a) The seismic signal contains information about two distinct sources during the eruption. (Top) A horizontal long-period source inverted from the surface waves [*Kanamori and Given*, 1982] and (bottom) shorter period vertical pulses. (b) A magnified plot of the vertical pulses. These pulses are the sources of the seismic waves in Figure 5.2 [*Kanamori et al.*, 1984].

the S waves indicate that the inclination of the source must be at least 60° from the horizontal. This estimate is further refined by observing that the dominant period of the S waves in Figure 5.2 is too long to be generated by the P wave source. If the P wave source generated any S waves, they are obscured by the longer period signal, that is, they are smaller amplitude than the traces in Figure 5.2b. Therefore, the inclination of the source must be much greater than 60° [*Kanamori et al.*, 1984]. We emphasize that the vertical orientation of the force is constrained by the azimuthal symmetry of the radiation patttern, rather than merely by the fact that the waves are observed on the vertical component.

The magnitudes of the single force sources are estimated using the amplitudes of the far-field P waves. The resulting source function in Figure 5.1b is two distinct

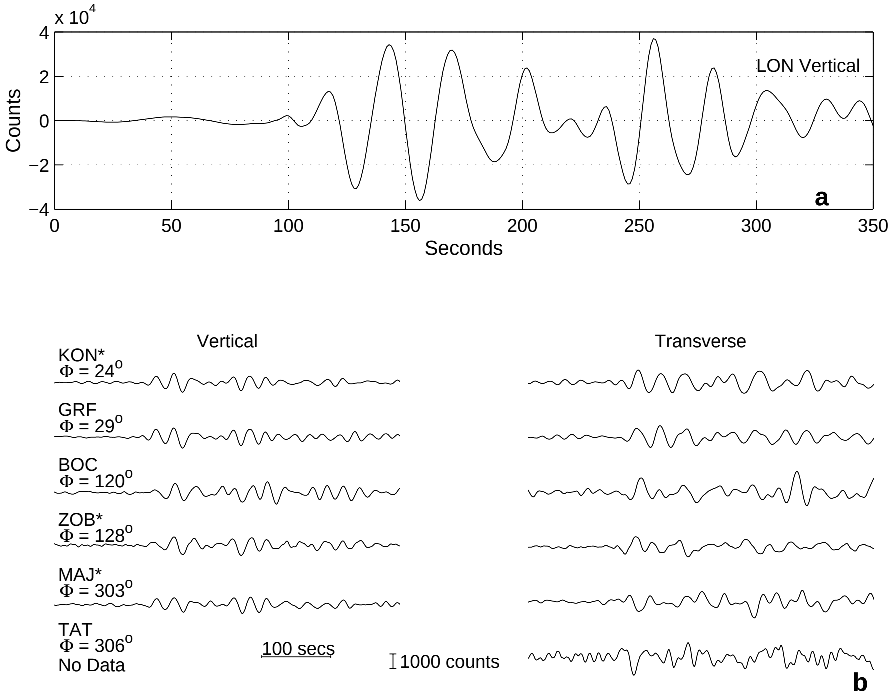

Figure 5.2 Body wave data for May 18, 1980 initial eruptive events. (a) Near-field vertical record from the WWSSN station at Longmire (LON). The distance Δ from the source was 67 km and the azimuth Φ was 26^o. (b) Far-field recordings of the P and S arrivals at Seismic Research Observatory Network (SRO) and Abbreviated Seismic Research Observatory Network (ASRO) stations arranged according to azimuths, Φ. P and S waves are observable on the vertical and transverse components, respectively, and the traces are aligned on the Jeffreys-Bullen arrival times of an event beginning at 1532:11. All amplitudes are normalized for geometric spreading to the distance of GRF, $\Delta = 76.2^o$. The instrument at TAT had the polarization of the horizontal components reversed in May 1980. Therefore, the observed trace is inverted in this figure. ASRO stations are denoted by asterisks. Their amplitudes are reduced by 2 and the traces are shifted 4 s to the left for comparison with the SRO stations. At a period of 25 s, an amplitude of 1000 counts in this figure corresponds to 0.2 μm of ground motion.

sequences of 20 to 30 s pulses with vertical component amplitudes of the order of 10^{12} N. The geometric constraints require that any horizontal components of these pulses have amplitudes less than half the vertical values. The largest pulse of the first sequence peaks at ~1532:45 which is the time that a sudden increase in the ash cloud height occurs in the photos of the eruption. Further support for this source inversion is provided by more detailed analysis of the seismic data including a full consideration of the radiation pattern [*Kanamori et al.*, 1984; *Burger and Langston*, 1985]. Similar lines of reasoning for the surface wave data yield a separate long-period (300 s) horizontal source that has been associated with traction at the base of the landslide [*Kanamori and Given*, 1982]. No direct measure of the horizontal component of the 20 to 30 s pulses is possible because the larger, longer period landslide dominated the transverse component signal.

Estimates of the erupted mass from both the geological and seismological observations are used in this paper to link the two types of data. The mass discharged at each stage of the eruption can be estimated by calculating the total mass of the corresponding geological deposits while assuming that the mass of the associated volatiles is negligible. Although gas was volumetrically and dynamically important, it is unlikely that its mass exceeded a few weight percent when the mixture was originally erupted from the vent [*Eichelberger and Hayes*, 1982; *Kieffer*, 1981]. The total blast deposit (unit A), which includes both country rock and juvenile dacite, had a volume of about 0.19 km^3 [*Moore and Sisson*, 1981]. This value is probably an overestimate of the material erupted from the vent since significant erosion and entrainment may have occurred as the blast cloud traveled down the slope of the mountain [*Kieffer and Sturtevant*, 1988]. The average density has been estimated as between 1660 [*Hoblitt et al.*, 1981] and 2180 kg/m^3 [*Glicken*, 1996]. The mass of the blast is therefore estimated to be 3.2×10^{11}–4.1×10^{11} kg. This value can be put into perspective by noting that the total mass of the juvenile magmatic products erupted on May 18, 1980 at Mount St. Helens is 5.2–7.0×10^{11} kg [*Christiansen and Peterson*, 1981; *Friedman et al.*, 1981].

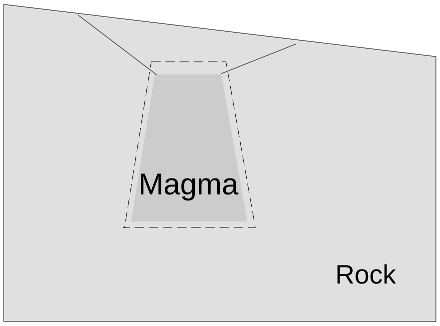

Figure 5.3 Cartoon of a control volume (dashed line) around a magma body (dark region).

Force balance

In order to understand the forces generating the seismic waves, we consider the balance of forces on the magma. The relevant dynamics are analyzed by modeling the magma as a fluid that is accelerated out of the ground by the depressurization that accompanied the landslide. The landslide removed the lid from a body of fluid under high pressure, and therefore the fluid was rapidly ejected.

We define a control volume around the magma in the ground (Figure 5.3). The geometry is simplified by representing the magma body as a trapezoidal volume. Each surface represents the area on which the resolved forces in the indicated direction act, rather than the literal wall of the magma body. The sum of the forces on the fluid equals the rate of change in the momentum of the fluid according to Newton's second law. The stresses acting on the fluid are gravitational and the pressures from the interactions with the atmosphere and the solid earth. The momentum can change in the fluid volume V either due to changes in the velocities internally or mass (and hence momentum) exiting through the surfaces of the volume. The sum of the forces exerted on the fluid $\sum F_{fl}$ is

$$\sum F_{fl} = -Mg + \sum F_{at} + \sum F_{se} = \frac{d}{dt}\int_V \rho u dV - \dot{M}_{fl} v \tag{5.1}$$

where M is the mass of the fluid, g is the magnitude of the gravitational acceleration, and the forces acting on the surface of the control volume are F_{at} from the atmosphere (including the part of the jet which is above the ground) and F_{se} from the solid earth. The velocity of the fluid exiting the control volume is v, u represents the velocity inside the control volume and ρ is the density of the fluid. The rate at which the mass of the fluid inside the control volume changes is $\dot{M}_{fl}$. Note that $\dot{M}_{fl} = -\dot{M}$ as the mass emission rate is positive when the mass in the control volume is decreasing. Forces are positive upwards.

Transient forces from the eruption that are coupled to the solid earth generate seismic waves. It is by observing these waves that we are able to determine the source parameters of the eruption so the "seismically observed force" F will be defined as

$$F \equiv -\sum F_{se}. \tag{5.2}$$

In this paper, we are interested in analyzing the observed vertical short-period seismic forces since the horizontal ones are not observed for the Mount St. Helens blast. Therefore, we will only consider the vertical force balance. Note that the only explicitly time dependent term in (5.1) is the time derivative of the momentum inside the control volume. We neglect this change of internal momentum, to derive a quasi-static approximation to (5.1),

$$\sum F_{fl} = \dot{M} v. \tag{5.3}$$

A full discussion of the validity of the quasi-static approximation for eruptive events is deferred until after the calculation.

Equation (5.3) is often referred to as the rocket equation [*Thompson*, 1972] since the sum of forces propelling a rocket upwards is equal to the momentum discharge rate of the fuel behind it. The blast at Mount St. Helens can be viewed as an inverted rocket where the momentum of the mass thrust into the air balances a net downward force into the ground.

The atmospheric force term $\sum F_{at}$ accounts for the fact that a supersonic jet can be "underexpanded," that is, at a pressure significantly above atmospheric at the vent. The net force of the jet on the fluid in the chamber is then the pressure in the jet P multiplied by the area of the vent. The sum of the atmospheric forces during the eruption is therefore

$$\sum F_{at} = -PA, \tag{5.4}$$

where A is the area of the vent. F can now be derived in terms of dynamic quantities by combining (5.1), (5.2), (5.3), and (5.4),

$$F = -\left(Mg + \dot{M}v + PA\right). \tag{5.5}$$

Equation (5.5) is a quasi-steady description of the force balance. It relates the observed force to the instantaneous dynamics of the jet but gives no insight into how this force varies over time. In order to understand the shape of the observed seismic pulses it is useful to reconsider the forces in terms of the geometry during depressurization. An eruptive pulse is divided into four representative stages in Figure 5.4.

Before the eruption (stage I) the magma is pressurized and at rest. We define the magnitude of the force exerted by the country rock at the top of the magma body as F_L. This is the force exerted by the lid prior to eruption and includes both the strength and weight of the caprock. The magnitude of the force on the bottom of the magma is denoted by F_b. F_b and F_L are defined to be the absolute values of the forces, and we explicitly note their signs in the following equations. For the purposes of the qualitative discussion of the time history of the force in this section, the vertical forces on the sides of the control volume are neglected. The difference between F_L and F_b is the net vertical force of the magma on the solid earth, i.e., the seismic force F.

$$F = F_L - F_b. \tag{5.6}$$

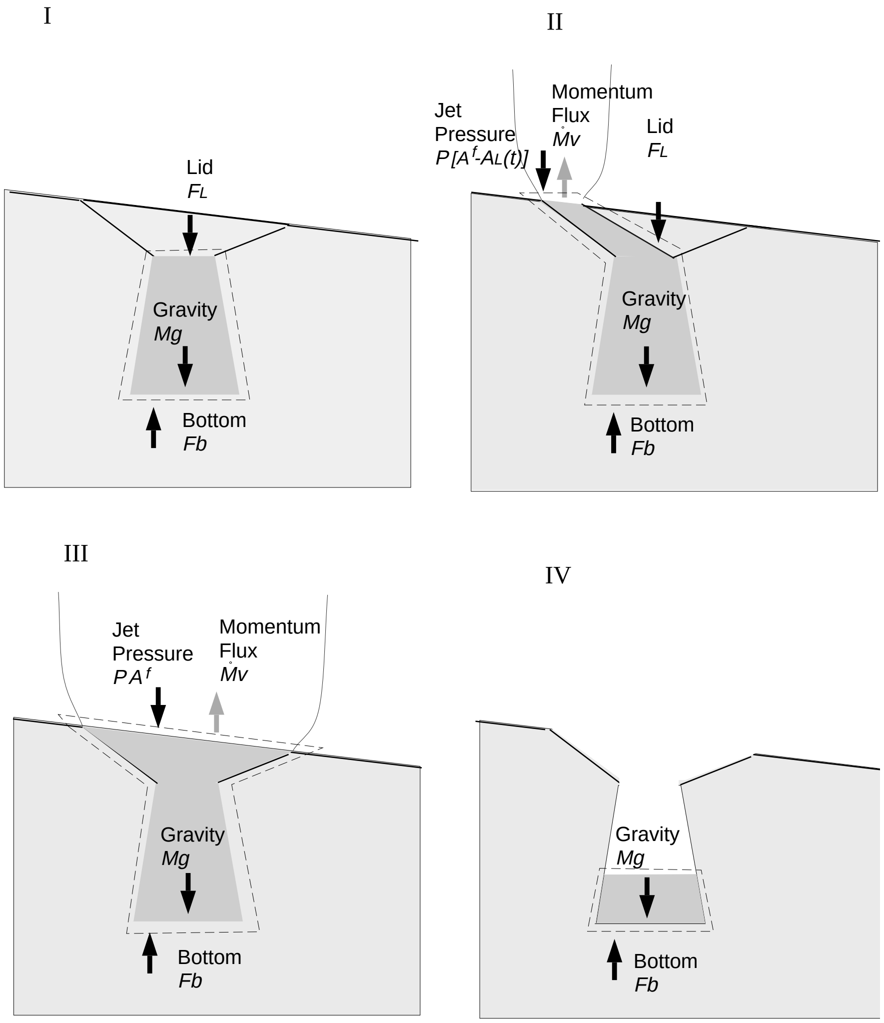

Figure 5.4 Cartoon of a control volume (dashed line) around a magma body (dark region) and the forces between the solid earth (lightly shaded area) and the magma acting during a single eruptive pulse. (I) The initial conditions. (II) The lid is removed in the landslide. (III) The fully developed jet. (IV) The final conditions.

During stage II the lid is removed in the landslide, the restraining force F_L decreases in time, and a jet forms. When the lid is removed, the pressure at the bottom of the magma body does not begin to fall until after a time τ required for the pressure drop, or rarefaction wave, to travel to the bottom of the magma body (Figure 5.5). During this time the lid continues to be accelerated. By the time the rarefaction wave reaches the bottom of the chamber, the pressure is dropping even faster at the top. Therefore, the downward force increases in magnitude throughout the lid removal process (stage II). Stage III in Figure 5.4 is when the lid is completely removed and the jet is fully developed. As the pressure in the chamber is released, the jet wanes and the magnitude of F decreases. The final phase, stage IV, occurs after the eruption. Only the weight of the remaining magma exerts a force on the ground. During the next pulse the process is repeated. We envision the initiation of the Mount St. Helens eruption to be a series of such depressurization events as the cap to the cryptodome broke into pieces. Each pulse released a section of the cryptodome that was separated from the other sections either by country rock or unvesiculated zones. We speculate that these separate regions may correspond to the heterogeneously vesiculated zones described by *Hoblitt and Harmons* [1993].

The geometric and dynamic considerations can be summarized by the following four equations:

$$\text{Stage I} \qquad F = \quad F_L^0 - F_b^0 \quad = -M_0 g \tag{5.7a}$$

$$\text{Stage II} \qquad F = \quad F_L(t) - F_b(t) \quad = -\{\dot{M}v + P\left[A^f - A_L(t)\right] + M(t)g\} \tag{5.7b}$$

$$\text{Stage III} \qquad F = \quad -F_b(t) \quad = -[\dot{M}v + PA^f + M(t)g] \tag{5.7c}$$

$$\text{Stage IV} \qquad F = \quad -F_b^f \quad = -M_f g \tag{5.7d}$$

where F_L^0, F_b^0, and M_0 are the initial lid force, bottom force, and mass, respectively,

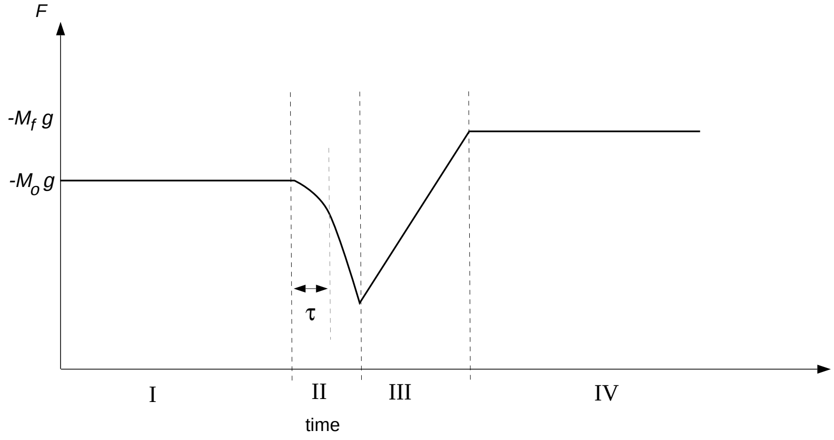

Figure 5.5 Schematic of seismic force F corresponding to the stages of Figure 5.4. The initial force is the preeruptive weight of the magma. Since a seismograph only records changes in F, this baseline value is arbitrary.

and F_b^f and M_f are the final bottom force and mass of the magma. A^f is the final area of the vent and A_L is the area that is covered by the lid. The force history of the entire process is shown schematically in Figure 5.5.

According to the above theory, the seismic force should be a function of the weight of the magma erupted, the thrust of the blast and the pressure in the jet. However, the seismological measurements are bandwidth limited. For the long, complex event on May 18, 1980, the gravitational unloading occurred over the entire 100 second series of pulses. In the far-field, such a long-period signal was obscured by the instrumental and ground noise. In the near-field (LON), the short record available shows spurious oscillations when the instrument is deconvolved [*Kanamori et al.*, 1984] and these artifacts obscure any long-period signal. Therefore, we neglected the gravitational terms Mg in (5.7a)–(5.7d) when analyzing the instrumentally recorded data for this event even though the magnitude of the gravitational unloading at the source may be as large as the other terms. For the same reason, the small vertical component of the longer period landslide force is neglected in the momentum balance. Fortunately, the photographic record documents that pulsations in the thrust and jet pressure occurred at somewhat shorter periods (∼20–30 s) [*Nielsen et al.*, 1989; *S. Malone*, Mount St. Helens timing, unpublished data, 1998]. These shorter period waves were recorded by the seismometers. If a higher frequency component of the unloading during the eruption was an important contribution to the seismic waves, the neglect of the upward force would mean that the values given for the downward thrusts here are underestimates.

We have now related the observed seismic force F to the momentum term $\dot{M}v$ and the jet force PA at each stage of the eruption and obtained a qualitative sense of the time variation of these variables. The only assumption about the fluid dynamics is that internal momentum changes are negligible. In order to proceed further, we need to relate the jet force PA to the momentum discharge rate $\dot{M}v$ which requires a specific model of the flow.

Flow model

We model the flow as an ideal gas expanding isentropically with a specified mass fraction ϕ of suspended particles. We follow *Kieffer* [1981] in modeling the fully developed flow geometry as a nozzle in which the discharge is a supersonic jet. For flows like the blast of Mount St. Helens that are volumetrically primarily gas, the equation of state of the mixture can be derived by describing the fluid as a mixture of two gases, one of which (the solid phase) is incompressible. Such a fluid is known as a dusty gas, or pseudogas, and is used as a model fluid in engineering applications ranging from solid fuel rocket engines to pipeline technology [*Rudinger*, 1980]. This formulation allows us to invoke a number of standard results from fluid mechanics. In particular, we can relate the flow parameters at the exit to the local sound speed. The ratio of the pressure term to the momentum discharge rate is then simply a function of the Mach number.

The equation of state of an equilibrium solid-gas mixture with noninteracting particles is [*Rudinger*, 1980]

$$P = \rho R_M T/(1-\xi) \tag{5.8}$$

where R_M is an ideal gas constant for the mixture, ρ, T and P are the density, temperature and pressure in the mixture, respectively, and ξ is the volume fraction of solids. The term "equilibrium" in this context means that the solid and gas phases flow at the same velocity (Appendix A).

The solid volume fraction is related to the mass fraction ϕ by

$$\xi = \phi \frac{\rho}{\rho_s}, \tag{5.9}$$

where ρ_s is the density of the solid. This equation of state is an ideal gas modified to account for the volume of an incompressible solid phase. In the limit of $\xi \to 0$ the fluid becomes an ideal gas as modeled by *Kieffer* [1981]. The isentropic sound speed c corresponding to (5.8) is [*Rudinger*, 1980]

$$c^2 \equiv \left(\frac{\partial P}{\partial \rho}\right)_s = \frac{\Gamma R_M T}{(1-\xi)^2} = \frac{\Gamma P}{\rho(1-\xi)}, \tag{5.10}$$

where Γ is defined to be the ratio of the specific heats for the mixture and is analogous to the adiabatic constant for an ideal gas. The specific heat of the mixture is the weighted average of the specific heats of the gas at constant pressure c_p and volume c_v and the specific heat of the solid c_s.

$$\Gamma = \frac{(1-\phi)c_p + \phi c_s}{(1-\phi)c_v + \phi c_s}. \tag{5.11}$$

The solid mass fraction ϕ is over 90%. The maximum whole rock H_2O content proposed for the preeruptive magma is 3–5 wt% based on both petrological and isotopic evidence [*Hoblitt and Harmons*, 1993]. A fraction of the H_2O remained trapped in vesicles and therefore dynamically inactive. Therefore, the minimum value of ϕ is 95%. The solid mass fraction ϕ remains constant throughout the flow of an equilibrium dusty gas even though the solid volume fraction ξ decreases significantly [*Rudinger*, 1980]. These high values of ϕ cause the adiabatic constant Γ to be nearly 1. In the limit of $\Gamma \rightarrow 1$, the nondissipative (isentropic) expansion of the fluid through a nozzle is isothermal. Rewriting (5.10),

$$PA = (1 - \xi)\rho c^2 A \quad ; \quad \Gamma = 1. \tag{5.12}$$

The momentum discharge rate at the exit, that is, the flux through the top surface of the control volume multiplied by the area, is

$$\dot{M}v = \rho v^2 A. \tag{5.13}$$

Therefore, the ratio of the pressure term to the momentum discharge rate is

$$\frac{PA}{\dot{M}v} = \frac{(1 - \xi)c^2}{v^2} = \frac{(1 - \xi)}{\mathcal{M}^2} \tag{5.14}$$

where $\mathcal{M}$ is the exit Mach number defined as the exit velocity v divided by the local sound speed c. Since $1 \geq \xi \geq 0$,

$$0 \leq PA \leq \dot{M}v\frac{1}{\mathcal{M}^2}. \tag{5.15}$$

Until now we have treated only the mass discharge rate of the bulk fluid. When we compare our results with the geological deposits, we will want to limit ourselves to the mass discharge rate of the solid phase. According to the definition of mass fraction ϕ, the solid mass discharge rate is simply $\phi\dot{M}$. As already noted, the mass fraction is nearly unity for the blast, and approximating the solid mass discharge rate as the total mass discharge rate is reasonable.

Equation (5.15) relates the pressure at the top of the control volume to the mass discharge rate. We combine (5.7b), (5.7c), and (5.15) with gravity neglected as discussed above and incorporate the definition of Mach number $\mathcal{M} \equiv v/c$, to obtain the inequality

$$\dot{M}c\mathcal{M} \le -F \le \dot{M}c\mathcal{M} + \frac{\dot{M}c}{\mathcal{M}}. \tag{5.16}$$

Since the seismically observed quantity is the force F, the inequality can be more usefully written as

$$\frac{-F}{c\mathcal{M}} \ge \dot{M} \ge \frac{-F}{c\mathcal{M}(1+1/\mathcal{M}^2)}. \tag{5.17}$$

If we knew the exit Mach number $\mathcal{M}$ and sound velocity c as a function of time as well as the force history exactly, we could estimate the mass discharge rate within a factor of $(1+1/\mathcal{M}^2)$. Unfortunately, it is not possible to know the precise values of these parameters for real eruptions, but we can make reasonable estimates for well-documented cases like Mount St. Helens using the considerations below.

Estimation of parameters

The force F as a function of time is approximately known from seismic data as detailed in the force balance section. We must estimate the Mach number $\mathcal{M}$ and the sound velocity c in order to utilize the bounds imposed by (5.17).

The Mach number $\mathcal{M}$ at the top of the control volume is computed by estimating the expansion in the crater (Appendix B). The area of the supersonic flow was probably smallest at the bottom of the crater. This point of minimum cross-sectional area is termed the "throat" in gas dynamics, and quantities measured there are denoted by a star subscript. The ratio of the area at the top of the crater to the area of the vent at the bottom of the crater walls as measured on a geological map [*Lipman*, 1981] is a rough measure of the expansion. The crater geometry during the blast was certainly different than the posteruptive topography since the mountain was undoubtedly changed during the 9-hour long violent eruption, so we consider a wide range of possible expansion ratios A/A_*. We estimate that the area of the jet increases by a factor between 4 and 40 which corresponds to Mach numbers of 2.1–3.4 assuming that the solid volume fraction at the throat ξ_* is between 0 and 0.5 (Appendix B). Since these values are based on approximate estimates, we will round the numbers and consider a range of $\mathcal{M}$ from 2 to 3.5 in this work. Note that this estimate of $\mathcal{M}$ is for the fully open vent of stage III. During stage II the jet expands less and the exit Mach number is smaller. Using values of $\mathcal{M}$ from 2 to 3.5 for both stages therefore results in a conservative estimate of the mass discharged.

The sound velocity c is estimated by rewriting (5.10) in terms of temperature [*Marble*, 1970; *Rudinger*, 1980]

$$c = \sqrt{\frac{\Gamma(1-\phi)RT}{(1-\xi)^2}}, \tag{5.18}$$

where R is the ideal gas constant for the volatile. We assume the gas phase was primarily H_2O with $R = 461.5$ J/kg /K. As can be seen from (5.18), the sound velocity is weakly dependent on the temperature and the volatile content of the erupting column. *Kieffer* [1981] proposed a cool reservoir of 600 K, 3.9 wt % H_2O and a density at the throat corresponding to a solid volume fraction ξ_* of about 30%. The expansion ratios A/A_* considered above reduce ξ to at most 4% at the exit and the corresponding sound velocity is ∼110 m/s. If Kieffer's model is adjusted to magmatic temperatures, that is, T=1200 K, the sound velocity is ∼150 m/s. *Eichelberger and Hayes* [1982] prefer a hotter, drier mixture of 1173 K, 0.7 to 1.7 wt % H_2O and at least 67 vol % solid throughout the flow. The volume fraction estimate of *Eichelberger and Hayes* [1982] is the fraction of H_2O trapped in the vesicles of pumice, rather than the dynamically active gas outside the clasts and therefore is an inappropriate value for the fluid fraction of the dusty gas. Individual pumice fragments were entrained in a gas-rich flow that is being modeled here. We consider a range of possible values of c from 100 to 150 m/s.

The values of $\mathcal{M}$ and c considered here correspond to exit velocities of 200–525 m/s. The exit velocity of the fluid at the top of the crater is poorly constrained by observations. The position of the front of the lateral blast has been analyzed in photos taken by G. Rosenquist [*Voight*, 1981]. Depending on the choice of timing schemes of the photos, the lateral blast front is estimated to have moved northward at ∼50–100 m/s within 1 km of the vent. However, the horizontal velocity of the blast downhill of the vent provides no constraint on the vertical velocity at the vent. Even if the blast were ejected in the direction of final motion, *Kieffer and Sturtevant* [1984] showed that front velocity is smaller than the fluid velocity at the exit.

The parameters above allow us to invert the force F for $\dot{M}$. For F in units of Newtons and $\dot{M}$ in kilogram per second, the numerical equivalent to (5.17) using the maximum and minimum values of $\mathcal{M}$ and c is

$$\frac{-F}{200} \geq \dot{M} \geq \frac{-F}{570}. \tag{5.19}$$

Figure 5.6 shows the total mass erupted computed from (5.19) using the observed forces and integrating the resulting $\dot{M}$ over time. The mass of the blast deposit as estimated from the geological evidence is also plotted as a dashed line for comparison. Around 1.6–4.6$\times 10^{11}$ kg of material was discharged vertically over the first 100 s. An additional 1–2.8$\times 10^{11}$ kg corresponds to the second set of forces.

Discussion

At this point in our study it is appropriate to ask how well-constrained is the force-time history by the seismic data. What errors are introduced in determining this function? Limited bandwidth seismic data does not tightly constrain the magnitude of the equivalent forces. However, the integrated force over time, or impulse, is robustly determined. Since the total mass erupted is a function of the impulse, the errors from the seismic data are negligible in comparison to the errors in estimating the parameters $\mathcal{M}$ and c. We can verify this claim by repeating the calculation of Figure 5.6 using an alternative inversion for the force-time history. The force-time function in Figure 5.7a is inverted from the far-field waveforms rather than the near-field data that was used for Figures 5.1b and 5.6a [*Kanamori et al.*, 1984]. Figure 5.7 shows that the impulse for the complete event, and hence the total erupted mass, is the same for these two models.

We can also quantify the error implicit in the quasi-static approximation. The in-

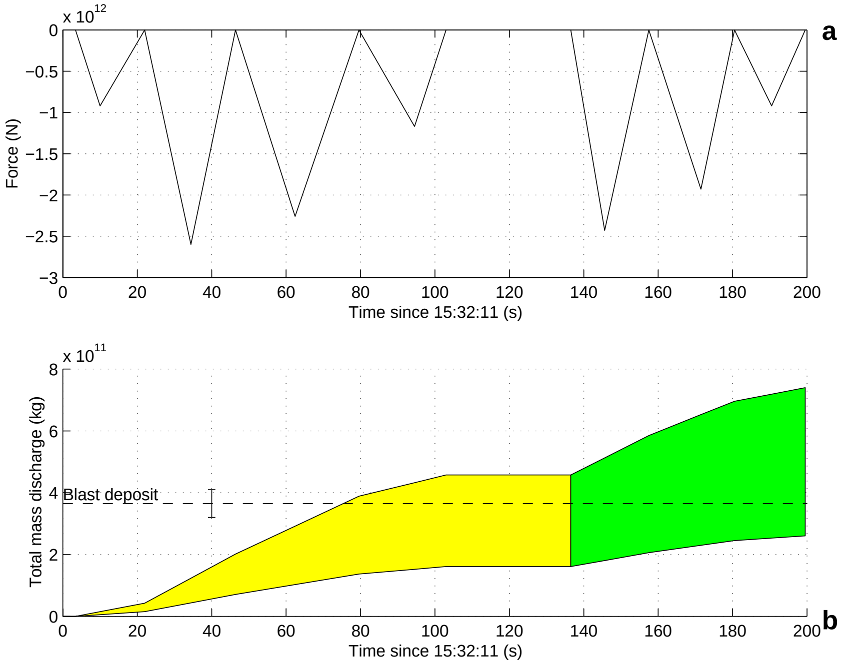

Figure 5.6 (a) The forces inverted from the seismic data of *Kanamori et al.* [1984]. (b) Estimate of mass erupted using data in Figure 5.6a and (5.17). The upper bound of the shaded region corresponds to c=100 m/s, $\mathcal{M}$=2 and the upper bound of (5.17); the lower bound corresponds to c=150 m/s, $\mathcal{M}$=3.5 and the lower bound of (5.17). The mass calculated from the second set of pulses is shaded darker because these later explosive events are not conclusively related to the blast (see text).

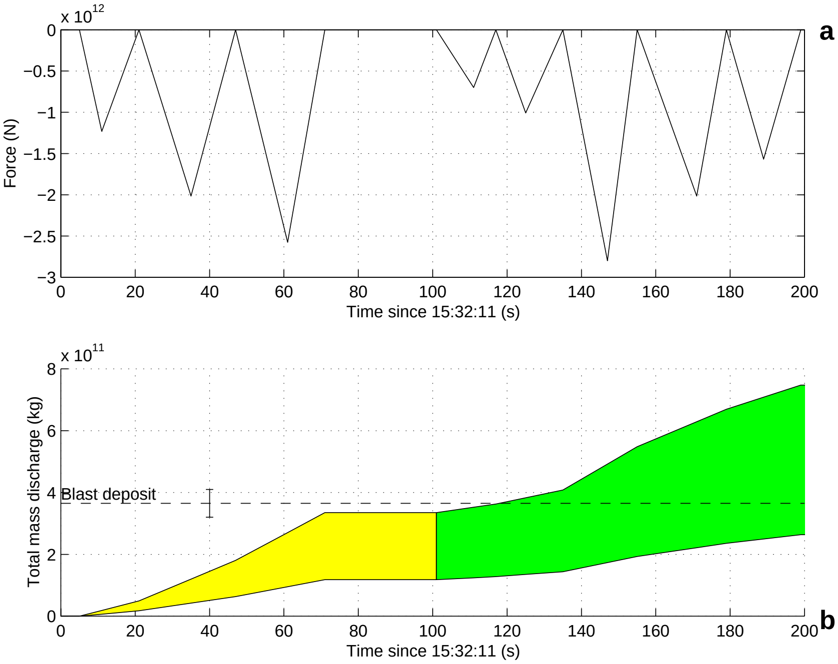

Figure 5.7 (a) An alternative force model from the seismic data of *Kanamori and Given* [1982]. (b) Estimate of mass erupted using data in Figure 5.7a and (5.17). Bounds are the same as in Figure 5.6. Note that the total mass at 200 s calculated here is the same as in Figure 5.6 since the total impulse is well constrained by the seismic data.

ternal momentum term acts as a storage term in the equations. When the boundaries are changed, e.g., the lid is removed, the internal momentum of the control volume is increased. The pressure drop travels through the body as a rarefaction wave. When it reaches the opposite boundary, momentum is transferred from inside the control volume to the outside by exerting a force on the bottom boundary. We can quantify the effect of internal momentum changes on the calculated total mass M by integrating the force balance equation over time from the beginning of the event t_0 to the end of the event t_f.

$$\int_{t_0}^{t_f} PA\,dt - \int_{t_0}^{t_f} F\,dt = \int_V \rho u\,dV\bigg|_{t_0}^{t_f} + M\bar{v}, \tag{5.20}$$

where $\bar{v}$ is an average exit velocity defined by

$$\bar{v} \equiv \frac{\int \dot{M} v\,dt}{M}. \tag{5.21}$$

If we assume that the internal momentum before (stage I) and after the event (stage IV) are both negligible, then the net effect of the internal momentum term is negligible. The total mass erupted M is robustly determined by neglecting internal momentum even though there may be some error in $\dot{M}$ as a function of time.

The initial hypothesis that the forces can be modeled as a series of jets must also be reexamined before the calculated results can be interpreted. How applicable is the cartoon of Figure 5.4 to each of the observed pulses? Some alternative processes are considered below.

The landslide undoubtedly was not a strictly horizontal phenomenon. How do we know that the pulses we observe are not simply responses to irregularities in the landslide movement? The landslide radiated seismic energy by shearing the ground at its base. The equivalent forces on the elastic medium were single forces parallel to the ground. Even if the landslide moved over steep topography in places, an equivalent force at an angle >60° to the horizontal is highly unlikely to be caused by the landslide at any time.

Another objection that might be raised is that the decreasing train from 22 to 103 seconds might be reflections due to only one initial explosive event. However, a simple seismic reflection from the Moho or some other geological feature would have occurred significantly faster than the observed 25–30 s between peaks. An atmospheric reflection would have been slower than the observed period by at least a factor of three. Another possibility is that the pulses are reverberating depressurization (rarefaction) waves in the magma body. It has been proposed that the main magma body was at approximately 7–9 km below the cryptodome [*Scandone and Malone*, 1985]. If the waves traveled at the sound velocity through the unfragmented magmatic conduit connecting the upper and lower chambers, the two-way travel time might be the requisite 25–30 secs. However, these reflections would result in further rarefaction waves and therefore a continually decreasing force rather than the observed series of triangular pulses. The increasing stage of the downward force (stage II) requires the removal of a lid. Each pulse must be a separate explosive event with attendant lid removal.

The coincidence of the first set of thrusts over the first 2 min with the photographically documented expansion suggests that the thrusts arose from the depressurization of sections of the cryptodome. It has been suggested that the second set of pulses from 1534.6 to 1535.3 UT is the result of a distinctly different process. *Moore and Rice* [1984] observed that at approximately the time of these forces a plume rose from the Spirit Lake or Toutle River area 8–10 km north of the vent. The most sustained and intense infrared radiation was also observed during 1534.4–1534.7 [*Moore and Rice*, 1984]. *Moore and Rice* [1984] interpret these events as either the unroofing of previously undisturbed cryptodome material in the landslide blocks or explosions produced by the interaction of the hot material with the water of the Toutle River drainage. Either interpretation of *Moore and Rice* [1984] would imply that the momentum balance would be somewhat different than in the model presented in the Force Balance section and the appropriate velocities would be unconstrained. However, the secondary explosion hypothesis of *Moore and Rice* [1984] is controversial. *Hoblitt* [1989] proposes that the late explosions accompanied the removal of slide block III, the former top of Mount St. Helens. In this scenario, our original analysis remains applicable. The secondary unroofing event occurred 2 min after the start of

Table 5.1. Mass Comparison

	Mass 10^{11} kg	Mass % Blast Deposit	Mass % Total Erupted
Blast deposit	3.2–4.1	100	35–60
Total juvenile	5.2–7.0	125–220	60–100
Total erupted	7–9	170–280	100
Computed mass a	1.6–4.6	40–90	20–65
Computed mass b	2.6–7.4	65–230	30–105

A comparison of the mass calculated for various stages of the eruption by various methods. Blast deposit mass is the mass of Unit A as discussed in the text. Total juvenile is the sum of the masses of magmatic products from all phases of the May 18 eruption [*Christiansen and Peterson*, 1981; *Friedman et al.*, 1981]. Total erupted is the sum of masses from all phases of the eruption, including the non-juvenile component of the blast. Computed mass a and b are the seismologically determined, vertically ejected mass. Computed mass a corresponds to the first 120 seconds of data; computed mass b includes both sets of pulses in Figure 5.6.

the eruption and produced jets that were obscured by the expanded blast cloud. Recognizing the ambiguity of the field evidence, we include the mass from the secondary explosions in darker shading in Figure 5.6b but acknowledge that the relationship between this mass and the deposits is uncertain.

The total mass of the eruption on May 18 was 7–9$\times10^{11}$ kg (Table 5.1). The calculations of this study imply that between 1.6 and 4.6$\times10^{11}$ kg, that is, 20%–65% of the total mass of the eruption, was discharged vertically at a rate of $\sim$2–6$\times10^{9}$ kg/s within the first 2 min of the eruption. A second set of explosions is equivalent to an additional 1–2.8$\times10^{11}$ kg of material. Including the second set of explosions brings the vertically erupted mass total to 2.6–7.4$\times10^{11}$ kg.

Interpretation

The term “lateral blast,” which is usually used to refer to the initial explosive event of the Mount St. Helens eruption, implies a predominantly horizontal flow. We can use the mass calculation above to evaluate this implicit assumption of the blast's direction. Figure 5.6 indicates that at least 40%, or at least 65% if the second set of events is included, of the mass of the blast deposit was involved in the vertical thrust events. We acknowledge that there may be substantial errors in both the geological deposit measurements and the seismic model, but even if we allow for this uncertainty, a substantial fraction of the blast deposit mass was jetted vertically from the vent. We reemphasize that this result is independent of the assumed crater geometry. As discussed in the Observations section, the seismic data constrains the measured component of the force, and hence the measured mass, to be subvertical. If each of the vertical pulses represents a thrust, the mass balance considerations make it unreasonable to assume that all of the blast deposit is composed of material that was initially ejected horizontally. Much, if not most, of the blast was initially ejected "vertically," that is, at an angle much greater than 60° from the horizontal. The destructive "lateral" blast was either composed of only a fraction of the mass in the deposit or the vertical blasts were redirected by the geometry of the crater and the force of gravity. Either possibility has implications for the quantitative understanding of directed blasts and their deposits.

Conclusions

This case study of Mount St. Helens has shown that the thrust inversion technique is an effective tool for measuring mass discharge rate. The results are reasonable although direct verification is imposssible since the mass ejection rate for this eruption cannot be measured by any other method. The mass ejection history is computed from the seismic data using only two parameters: Mach number and sound velocity. We now return to the original motivations for measuring $\dot{M}$ mentioned in the introduction in order to evaluate the importance of our results for the case of Mount St. Helens. We have found a quantitative meaure of "explosivity"; over the first 100 s $\sim$2–6 $\times 10^9$ kg/s were ejected. The blast style of eruptions is indeed distinguished by extremely high mass ejection rates. Greater than 20% of the total erupted products of May 18, 1980, were ejected in less than 2 min of the blast. Much of the mass was concentrated into even shorter pulses at the peaks of the mass ejection events. Further quantification of both the eruptive style and the potential hazard is possible by noting that the "violence," that is, momentum [*Walker*, 1980], of the eruption is directly measured by the seismic force F to within a factor of $(1 + 1/\mathcal{M}^2)^{-1}$ (see equation 5.17). We propose an index of violence M_v defined as

$$M_v \equiv \frac{1}{2} \log |F_{\max}| , \qquad (5.22)$$

where $F_{\max}$ is the peak value of the seismic force in Newtons. The factor of 1/2 is included to produce values comparable to standard volcanic indices and earthquake magnitude scales. For the Mount St. Helens blast $M_v = 6.2$. Although the M_v index is well-defined by seismic parameters, its physical interpretation should be used with caution. Volcanic processes other than explosive ejection of mass can also produce single forces. Examples include caldera collapse and landslides. In these cases, the equivalent erupted mass interpretation of M_v would not be applicable.

We have linked the blast process to its products, but the relationship has proved problematic. Is all of the blast deposit material from the laterally expanding, devasting blast or is a significant fraction of deposition from the accompanying ash cloud? Alternatively, we can interpret the results as addressing the final motivation—constraints on dynamical models. In this case, the important problem raised by this study concerns the mechanism by which the blast was directed. Why was it lateral? These questions pose challenges for future research and direct our understanding of the mass ejection process. The formulation presented here is generally applicable to any rapid, explosive volcanic eruption with significant jetting and can be used to study dome collapses and strombolian events as well as directed blasts. A major weakness of this study is that direct verification of the results by an independent method is impossible. Future studies of smaller, confined eruptions may provide more constraints on the accuracy of the method.

Appendix A: Equilibrium flow conditions

If the gas and particle phases initially travel at different velocities, the relaxation time τ_v required for the particles to be accelerated to the gas velocity can be derived by balancing the particle acceleration against the Stokes drag. The standard result is

$$\tau_v = \frac{\rho_p D^2}{18\eta}, \tag{5.23}$$

where ρ_p is the density of the particles, D is the diameter of the particles, and η is the gas viscosity corrected for the presence of particles [*Rudinger*, 1980]. This formulation of τ_v overestimates the relaxation time for high Reynolds number flows and therefore provides a conservative estimate of the time for equilibration. Thermal equilibrium is obtained over roughly the same timescale [*Rudinger*, 1980]. Appropriate values for ρ_p and η are 2000 kg/m^3 [*Hoblitt and Harmons*, 1993] and 8×10^{-5} Pa s [*Rudinger*, 1980; *Sengers and Watson*, 1986]. The median particle size in the blast deposit is ~0.1–1 mm [*Moore and Sisson*, 1981]. Equation (5.23) implies that under the above conditions all particles 1 mm in diameter or less are in equilibrium with the gas phase after about 1.4 s. This value of τ_v is compared to the transit time in order to establish whether or not the flow is in equilibrium at the throat. All of the dusty gas results used in this chapter are applied downstream of the throat. If the flow is in equilbrium at the throat, an equilbrium dusty gas is a good approximation for the entire downstream region.

Mass balance considerations require the fragmentation wave to travel into the initially stationary cryptodome at a velocity v_f such that

$$v_f = -v_1 \frac{\rho_1}{\rho_2 - \rho_1}, \tag{5.24}$$

where ρ_1 and ρ_2 are the densities behind and in front of the fragmentation wave, respectively, and v_1 is the average velocity of the flow behind the fragmentation wave (Figure 5.8). The fragmentation wave travels the distance L_f from the throat in the time t_f. Therefore, a particle travelling from the fragmentation wave at the velocity

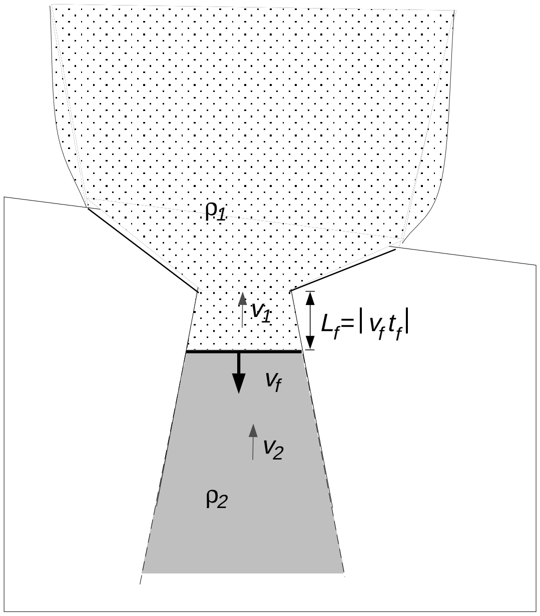

Figure 5.8 Fragmentation wave travelling into the magma body. The dark area is the magma and the stippled region is the fragmented dusty gas. Equilibrium flow is established at the throat when $\tau_v < L_f/v_1$.

v_1 reaches the throat at time $t_1 = L_f/v_1$. The transit time t_1 scales as

$$t_1 \sim t_f \rho_1/(\rho_2 - \rho_1). \tag{5.25}$$

The maximum value of ξ_* considered in this work is 0.5 which correponds to $\rho_1/\rho_2 = 0.5$ for $\phi = 1$. If $\xi_* = 0.5$, then $t_1 = t_f$ and equilbrium flow is established at the throat if $\tau_v < t_f$. We assume that the fragmentation wave starts from the throat at the beginning of the lid removal, i.e., at the beginning of stage II in Figure 5.4. Using the value for τ_v that is calculated above, the dusty gas is in equilbrium after the first 1.4 s of a pulse. This result is valid if the solid volume fraction is large. In this case, equilibrium flow is a justified simplification for most ($> 90\%$) of the pulsation process.

A dilute flow ($\rho_1/\rho_2 \to 0$) can have an arbirtrarily short transit time and may not reach equilbrium by the throat. In this case, it is useful to consider the situation with no coupling between the gas and the solid phases as a alternative end-member behavior. The exit velocity used in the momentum balance is the solid particle velocity and it must be estimated by a ballistic calculation. If the gas and solid phases are uncoupled, drag is negligible and the ballistic velocity v_b is

$$v_b = \sqrt{2gh}, \tag{5.26}$$

where g is the gravitational acceleration and h is the eruption height above the vent. In order to conservativly estimate $\dot{M}$, we only consider the maximum value of v_b corresponding to the maximum projectile height. The initial blast cloud was observed to rise to a maximum height of 5 or 6 km above the vent [*Sparks et al.*, 1986]. Therefore, the maximum ballistic exit velocity is between 310 and 340 m/s. The jet pressure is atmospheric since it is unlikely that a pure gas phase would reach the supersonic velocities. The equivalent expression to (5.19) in SI units is

$$\frac{-F}{310} \geq \dot{M} \geq \frac{-F}{340}. \tag{5.27}$$

These values are within the range considered in (5.19). Therefore, no separate calculations are required.

Appendix B: Estimating the Mach number

We can estimate the Mach number for a steady expanding isentropic dusty gas by first considering the conservation of momentum [*Rudinger*, 1980]. As the fluid flows in the x direction,

$$\rho u \frac{du}{dx} = -\frac{dP}{dx}, \tag{5.28}$$

where u is the velocity of the flow and P is the pressure. Eliminating the differential dx from (5.28) results in an expression for dP/du. The pressure of a fluid expanding isentropically is related to the density by the sound speed. Rearranging (5.28) and substituting the sound velocity $c^2 \equiv (\partial P/\partial \rho)_s$,

$$\frac{-u}{c^2} du = 1/\rho \, d\rho. \tag{5.29}$$

If the sound speed were constant, the integration of (5.29) would be trivial. However, a dusty gas generally has a variable solid volume fraction ξ as it expands. We substitute the sound speed from equation (5.10) into (5.29) and integrate from the throat to the vent to find a general expression for the Mach number with variable ξ in the limit of the specific heat ratio $\Gamma \to 1$.

$$\frac{1}{2}(1 - \mathcal{M}_*^2) = \frac{(1-\xi_*)^2}{1-\xi} - (1-\xi_*) + (1-\xi_*)^2 \ln \frac{\rho}{\rho_*} \frac{1-\xi_*}{1-\xi}. \tag{5.30}$$

The modified Mach number $\mathcal{M}_*$ denotes the velocity at the vent u divided by the sound velocity at the throat c_*. We can eliminate the density ratio ρ/ρ_* by noting that mass is conserved throughout the flow and $u_* = c_*$ [*Liepmann and Roshko*, 1957],

$$\frac{\rho}{\rho_*} = \frac{u_* A_*}{uA} = \frac{A_*}{A\mathcal{M}_*}. \tag{5.31}$$

Similarly, since no solid magmatic material transforms to the volatile phase or vice versa, the solid phase is also conserved

$$\frac{\xi}{\xi_*} = \frac{u_* A_*}{uA} = \frac{A_*}{A\mathcal{M}_*}. \tag{5.32}$$

Combining (5.30), (5.31) and (5.32),

$$\frac{1}{2}(1 - \mathcal{M}_*^2) = \frac{(1 - \xi_*)^2}{1 - \xi_* \frac{A_*}{A\mathcal{M}_*}} - (1 - \xi_*) + (1 - \xi_*)^2 \ln \frac{1 - \xi_*}{\frac{A\mathcal{M}_*}{A_*} - \xi_*}. \tag{5.33}$$

Given a volume fraction ξ_* and expansion ratio A/A_*, (5.33) can be solved numerically for $\mathcal{M}_*$. The modified Mach number $\mathcal{M}_*$ is related to $\mathcal{M}$ by the ratio of the sound speeds of the vent and the throat. In the limit of $\Gamma \to 1$ the fluid expansion is isothermal and the sound speed varies only due to the variations in solid volume fraction,

$$\mathcal{M} = \mathcal{M}_* \frac{1 - \xi}{1 - \xi_*} = \frac{\mathcal{M}_* - \xi_* \frac{A}{A_*}}{1 - \xi_*}. \tag{5.34}$$

Equations (5.33) and (5.34) together provide a complete method of estimating the Mach number for particular values of ξ_* and A/A_*. In the limit of negligible solid volume fraction ($\xi_* \to 0$), the sound speed remains constant and the Mach number can be calculated numerically with the simplified equation

$$\frac{A}{A_*}\mathcal{M} = \exp\left(\frac{\mathcal{M}^2 - 1}{2}\right). \tag{5.35}$$

References

Burger, R., and C. Langston, Source mechanism of the May 18, 1980, Mount St. Helens eruption from regional surface waves, *J. Geophys. Res.*, *90*, 7653–7664, 1985.

Christiansen, R. L., and D. W. Peterson, Chronology of the 1980 eruptive activity, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 17–30, 1981.

Criswell, C. W., Chronology and pyroclastic stratigraphy of the May 18, 1980, eruption of Mount St. Helens, *J. Geophys. Res.*, *92*, 10237–10266, 1987.

Eichelberger, J. C., and D. B. Hayes, Magmatic model for the Mount St. Helens blast of May 18, 1980, *J. Geophys. Res.*, *87*, 7727–7738, 1982.

Friedman, J. D., G. R. Olhoeft, G. R. Johnson, and D. Frank, Heat content and thermal energy of the June dacite dome in relation to total energy yield, May–October 1980, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 557–567, 1981.

Glicken, H., Rockslide-debris avalanche of the May 18, 1980, Mount St. Helens Volcano, Washington, *U.S. Geol. Surv. Open-file Rep.*, *96–677*, 1996.

Gorshkov, G. S., Directed volcanic blasts, *Bull. Volcan.*, *26*, 83–87, 1963.

Hickson, C., The May 18, 1980, eruption of the Mount St. Helens, Washington state: A synopsis of events and review of Phase I from an eyewitness perspective, *Geosci. Can.*, *17*, 127–131, 1990.

Hoblitt, R. P., Reply by R.P. Hoblitt to comment by M. Rosi on "May 18, 1980, Mt. St. Helens deposits in South Coldwater Creek, Washington," by R. V. Fisher, H. Z. Glicken, and R. P. Hoblitt, *Bull. Volcanol.*, *52*, 71–73, 1989.

Hoblitt, R. P., and R. S. Harmons, Bimodal density distribution of cryptodome dacite from the 1980 eruption of Mount St. Helens, Washington, *Bull. Volcanol.*, *55*, 421–437, 1993.

Hoblitt, R. P., and C. D. Miller, Comment on Walker & McBroome (1983), *Geology*, *12*, 692–693, 1984.

Hoblitt, R. P., C. D. Miller, and J. W. Vallance, Origin and stratigraphy of the deposit produced by the May 18 directed blast, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 401–420, 1981.

Kanamori, H., and J. W. Given, Analysis of long-period seismic waves excited by the May 18, 1980, eruption of Mount St. Helens - a terrestrial monopole?, *J. Geophys. Res.*, *87*, 5422–5432, 1982.

Kanamori, H., J. W. Given, and T. Lay, Analysis of seismic body waves excited by the Mount St. Helens eruption of May 18, 1980, *J. Geophys. Res.*, *89*, 1856–1866, 1984.

Kieffer, S. W., Fluid dynamics of the May 18 blast at Mount St. Helens, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 379–400, 1981.

Kieffer, S. W., and B. Sturtevant, Laboratory studies of volcanic jets, *J. Geophys. Res.*, *89*, 8253–8268, 1984.

Kieffer, S. W., and B. Sturtevant, Erosional furrows formed during the lateral blast at Mount St. Helens, May 18, 1980, *J. Geophys. Res.*, *93*, 14793–14816, 1988.

LaCroix, A., Remarques sur les materiaux de projection des volcans et sur la genese des roches pyroclastiques qu'ils constituent, *Centenaire de la Societe Geologique de Francais*, *2*, 431–472, 1930.

Liepmann, H. W., and A. Roshko, *Elements of Gasdynamics*, John Wiley & Sons, New York, 1957.

Lipman, P. W., Geologic map of deposits and features of 1980 eruptions of Mount St. Helens, Washington, *U.S. Geol. Surv. Prof. Pap.*, *1250*, plate 1, 1981.

Marble, F. E., Dynamics of dusty gases, *Ann. Rev. Fluid Mech.*, *2*, 397–446, 1970.

Moore, J. G., and C. J. Rice, *Explosive Volcanism: Inception, Evolution and Hazards*, chap. Chronology and Character of the May 18, 1980, Explosive Eruptions of Mount St. Helens, pp. 133–142, National Academy Press, Washington, D.C., 1984.

Moore, J. G., and T. W. Sisson, Deposits and effects of the May 18 pyroclastic surge, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 421–438, 1981.

Newhall, C., and S. Self, The volcanic explosivity index (VEI): An estimate of explosive magnitude for historical volcanism, *J. Geophys. Res.*, *87*, 1231–1238, 1982.

Nielsen, E., R. Waitt, and S. Malone, Eyewitness accounts and photographs of the 18 May 1980 eruption of Mount St. Helens, Wash., *New Mexico Bureau of Mines and Mineral Resources, Bulletin*, *131*, 203, 1989.

Nishimura, T., Source parameters of the volcanic eruption earthquakes at Mount Takachi, Hokkaido, Japan, and a magma ascending model, *J. Geophys. Res.*, *100*, 12,465–12,474, 1995.

Rosenbaum, J. G., and R. B. Waitt, Summary of eyewitness accounts of the May 18 eruption, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 53–67, 1981.

Rudinger, G., *Fundamentals of Gas-Particle Flow*, Elsevier Sci., New York, 1980.

Sarna-Wojicki, A., A. Shipley, R. B. Waitt, D. Dzurisin, and S. H. Wood, Areal distribution, thickness, mass, volume, and grain size of air-fall ash from the six major eruptions of 1980, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 577–600, 1981.

Scandone, R., and S. D. Malone, Magma supply, magma discharge and readjustments of the feeding system of Mount St. Helens during 1980, *J. Volcanol. Geotherm. Res.*, *23*, 239–262, 1985.

Sengers, J., and J. Watson, Improved international formulations for the viscosity and thermal-conductivity of water substance, *J. of Phys. Chem. Ref. Data*, *15*, 1291–1314, 1986.

Sisson, T. W., Blast ashfall deposit of May 18, 1980 at Mount St-Helens, Washington, *J. Volcanol. Geotherm. Res.*, *66*, 203–216, 1995.

Sparks, R. S. J., J. G. Moore, and C. J. Rice, The initial giant umbrella cloud of the May 18th, 1980, explosive eruption of Mount St. Helens, *J. Volcanol. Geotherm. Res.*, *28*, 257–274, 1986.

Sparks, R. S. J., M. I. Bursik, S. N. Carey, J. Gilbert, L. Glaze, H. Sigurdsson, and A. W. Woods, *Volcanic Plumes*, John Wiley, New York, 1997.

Thompson, P. A., *Compressible-Fluid Dynamics*, McGraw-Hill, New York, 1972.

Uhira, K., and M. Takeo, The source of explosive eruptions of Sakurajima Volcano, Japan, *J. Geophys. Res.*, *99*, 17,775–17,789, 1994.

Uhira, K., H. Yamasato, and M. Takeo, Source mechanism of seismic waves excited by pyroclastic flows observed at Unzen Volcano, Japan, *J. Geophys. Res.*, *99*, 17,757–17,773, 1994.

Voight, B., Time scale for the first moments of the May 18 eruption, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 69–86, 1981.

Waitt, R. B., Comment on Walker and McBroome (1983), *Geology*, *12*, 693, 1984.

Waitt, R. B., and D. Dzurisin, Proximal air-all deposits from the May 18 eruption: Stratigraphy and field sedimentology, *U.S. Geol. Surv. Prof. Pap.*, *1250*, 601–616, 1981.

Walker, G., The Taupo pumice: Produce of the most powerful known (ultraplinian) eruption?, *J. Volcanol. Geotherm. Res.*, *8*, 69–94, 1980.

Walker, G. P. L., and L. A. McBroome, Mount St. Helens 1980 and Mount Pelee 1902: flow or surge?, *Geology*, *11*, 571–574, 1983.

Wilson, L., R. S. J. Sparks, T. C. Huang, and N. D. Watkins, The control of volcanic column heights by eruption energetics and dynamics, *J. Geophys. Res.*, *83*, 1829–1836, 1978.

Wilson, L., R. S. J. Sparks, and G. P. L. Walker, Explosive volcanic eruptions; IV, the control of magma properties and conduit geometry on eruption column behaviour, *Geophys. J. R. Astr. Soc.*, *63*, 117–148, 1980.