Tom Heaton's Engineering Seismology class notes 10/8/2022

Chapter 8   Earthquake Scaling

Earthquake Intensity

Earthquakes occur in many different places in the Earth, some are shallow in the continental crust, others are along the interface between a subducting oceanic plate and a continental margin, and yet others are deep within the Earth (up to 650 km). Before the invention and deployment of seismographs, these earthquakes were largely described by their perceived maximum shaking intensity. There have been several shaking intensity scales developed through the years; the Rossi-Forel was defined in 1883 and it was the most common scale until the **Mercalli Intensity scale (MI)** was defined in 1902. This scale was modified in 1931 (intensities XI and XII were added to describe very damaging shaking and the definition of Modified **Mercalli Intensity scale (MMI)** is given in Table 8-1. Earthquakes are sometimes characterized by the maximum MMI reported. However, since large MMI values are based largely on observed damage to structures, maximum MMI is not a consistent way to characterize the size of earthquakes. In particular, the strongest shaking from an earthquake most commonly occurs in areas that do not have a sufficient density of buildings to characterize the intensity of the shaking. Furthermore, building construction changes dramatically with time and from region to region.

Modified Mercalli Intensity Scale

I. People do not feel any Earth movement.

II. A few people might notice movement if they are at rest and/or on the upper floors of tall buildings.

III. Many people indoors feel movement. Hanging objects swing back and forth. People outdoors might not realize that an earthquake is occurring.

IV. Most people indoors feel movement. Hanging objects swing. Dishes, windows, and doors rattle. The earthquake feels like a heavy truck hitting the walls. A few people outdoors may feel movement. Parked cars rock.

V. Almost everyone feels movement. Sleeping people are awakened. Doors swing open or close. Dishes are broken. Pictures on the wall move. Small objects move or are turned over. Trees might shake. Liquids might spill out of open containers.

VI. Everyone feels movement. People have trouble walking. Objects fall from shelves. Pictures fall off walls. Furniture moves. Plaster in walls might crack. Trees and bushes shake. Damage is slight in poorly built buildings. No structural damage.

VII. People have difficulty standing. Drivers feel their cars shaking. Some furniture breaks. Loose bricks fall from buildings. Damage is slight to moderate in well-built buildings; considerable in poorly built buildings.

VIII. Drivers have trouble steering. Houses that are not bolted down might shift on their foundations. Tall structures such as towers and chimneys might twist and fall. Well-built buildings suffer slight damage. Poorly built structures suffer severe damage. Tree branches break. Hillsides might crack if the ground is wet. Water levels in wells might change.

IX. Well-built buildings suffer considerable damage. Houses that are not bolted down move off their foundations. Some underground pipes are broken. The ground cracks. Reservoirs suffer serious damage.

X. Most buildings and their foundations are destroyed. Some bridges are destroyed. Dams are seriously damaged. Large landslides occur. Water is thrown on the banks of canals, rivers, lakes. The ground cracks in large areas. Railroad tracks are bent slightly.

XI. Most buildings collapse. Some bridges are destroyed. Large cracks appear in the ground. Underground pipelines are destroyed. Railroad tracks are badly bent.

XII. Almost everything is destroyed. Objects are thrown into the air. The ground moves in waves or ripples. Large amounts of rock may move.

Table 8.1. Definition of the Modified Mercalli Intensity Scale.
Full descriptions are from: Richter, C.F., 1958. *Elementary Seismology.* W.H. Freeman and Company, San Francisco, pp. 135-149; 650-653.

U.S. earthquakes that occurred from the 1920's into the 1970's were assigned MMI's by the U.S. Coast and Geodetic Survey (merged with the USGS in 1978). The USCGS conducted routine surveys to prepare MMI intensity maps that were published regularly in a series called U.S. Earthquakes. The practice of compiling these data was largely discontinued in the 1980's (too much manpower), and earthquakes were primarily described by other parameters (magnitude, pga, spectral acceleration). However, the widespread use of the Internet has allowed MMI compilations to be automatically managed through a USGS project known as "*did you feel it*?" https://earthquake.usgs.gov/data/dyfi/. In addition to shaking intensity observations, it has become common to characterize shaking intensity with observations of pga and pgv derived from strong-motion seismic data. In particular, relationships between reported MMI's and nearby recorded pga's and pgv's were used to derive quantitative relations that predict MMI from pga and pgv (Wald, D., V. Quitoriano, T. Heaton, and H. Kanamori, 1999, Relationships between peak ground acceleration, peak ground velocity, and Modified Mercalli Intensity in California, Earthquake Spectra, 15, 557-564 pdf).

This is now referred to as Instrumental Modified Mercalli Intensity (IMM), which is given by

$$IMM = 3.66Log(PGA) - 1.66 \text{ for V} \leq \text{IMM} \leq \text{VIII}$$
$$IMM = 3.66Log(PGV) - 1.66 \text{ for V} \leq \text{IMM} \leq IX$$
8.1

equation reference goes here

$$IMM = 2.2\log(pga) + 1 \qquad IMM \leq V \qquad\qquad 8.2$$

IMM uses the principle that smaller intensities are based on human perception of shaking. In these cases, humans seem to be better at discriminating pga than they are pgv. In contrast, higher levels of intensity are based on observations of structural damage; this seems to correlate best with pgv.

Japan has a long history of important earthquakes and a Japanese intensity scale has evolved through time. The Japan Meteorological Agency (JMA) intensity scale (aka, Shindo scale) is summarized in table 8.2

Intensity **0**	Imperceptible to people.	
		0.8% to 2.5% g
Intensity **1**	Some people in the building feel it.	
Intensity **2**	Many people in the building feel it. Some people awaken, if the quake strikes at night.	
Intensity **3**	Felt by most people in the building. Some people are frightened.	0.8% to 2.5% g
Intensity **4**	Many people are frightened. Some people try to escape from danger. Most people awaken, if the quake strikes at night.	2.5% to 8% g
Intensity **5**lower	Most people try to escape from danger. Some people find it difficult to move.	8% to 14% g
Intensity **5**upper	Many people are very frightened and find it difficult to move.	14% to 25% g
Intensity **6**lower	Difficult to keep standing.	25% to 32% g
Intensity **6**upper	Impossible to keep standing and to move without crawling.	32% to 40% g
Intensity **7**	Thrown around by the shaking. Impossible to move at will.	More than 40% g

Figure 8.1. JMA intensity scale

Shaking intensity scales measure ground shaking at a point, BUT there must be structures to characterize the damage, and locally intense shaking may not reflect the overall size of an earthquake. One method to measure overall size using intensity data has been to estimate the area that experienced at least a given level of intensity. In particular, the area that experienced MMI ≥ IV has been helpful in quantifying the size of earthquakes for which only intensity data is available (pre 20th century). Intensity IV is convenient, since it is small enough to occur over a large area while still large enough to be reported in newspapers. (Estimating Earthquake Location and Magnitude from Seismic Intensity Data, 1997, Bakun W. and C. M. Wentworth, Bulletin of the Seismological Society of America, Vol. 87, No. 6, pp. 1502-1521).

Earthquake Magnitude

Charles Richter is usually identified as the designer of the first practical earthquake magnitude scale, which is now referred to as M_L (or local magnitude). When a seismologist tells you the "Richter magnitude" of an earthquake, he or she often intend to mean "local magnitude." Unfortunately, the non-seismologists seems to think that earthquakes are only assigned Richter magnitudes, which is definitely not true. Richter's magnitude scale was designed in the 1930's to assign a relative size to earthquakes that were recorded on the seismic network operated by Caltech's Seismological Laboratory in southern California. This network was comprised of horizontal-component, optical, torsion, seismometers designed by Harry Wood (Caltech seismologist) and John Anderson (an astronomer with the Hale Observatory). This instrument has a natural period of 0.8 s.; 80% damping, a displacement transducer (it recorded light deflected by a mirror onto photographic paper on a rotating drum) with a maximum gain of 2,800 (see chapter 2). The definition of M_L
is

$$M_L \equiv \log\left(\frac{A}{A_0}\right) = \log A - \log A_0 \qquad\qquad 8.3$$

 Where A is the maximum amplitude (in mm) of a Wood-Anderson seismogram. The original definition only applies for a seismogram recorded at an epicentral distance of 100 km. Richter arbitrarily chose M = 3.0 to correspond to a peak seismogram amplitude of 1 mm. That is he defined that $A_0\left(\Delta = 100 \text{ km}\right) = 3.0$. Since earthquakes were recorded at a range of distances, Richter derived a distance correction factor , A_0 $\left(\Delta\right)$, to correct the amplitudes to the standard distance of 100 km (see Fig. 8.2).
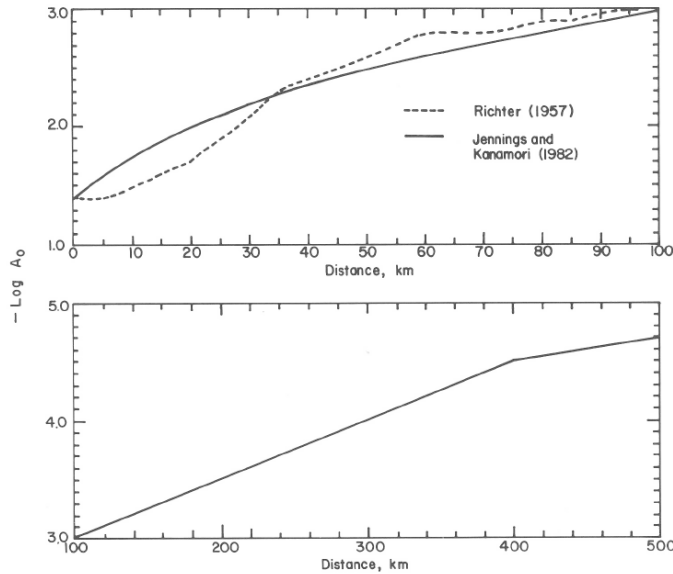
Fig. 8.2 Distance correction factor for Richter's local
magnitude scale, M_L. The solid line shows Richter's definition,
while the dotted line is from a later study by Kanamori. H. and
Jennings, P.C., 1978. Determination of local magnitude, M,
from strong-motion accelerograms. Bull. Seismol. Sot. Am.,
68: 471-485.

Since a Wood-Anderson seismometer is a linear single-degree-of-freedom oscillator, the
measured peak amplitude is identical to a 0.8 s, 80 % damped displacement-response
spectrum.   Since the damping is much larger than is typically used in structural
engineering, the amplitude used in M_L correlates well with pgv, which also seems to be
a good measure of shaking intensity; there is a detailed discussion of the relationship
between pgv and 80% damped response spectra in Chapter 6. . M_L was used from the
1920's through the 1980's as the standard magnitude to prepare catalogs of seismic
activity in California (it has been adopted in other countries as well).

There were some important limitations to the use of M_L; in particular, earthquakes larger
than M_L 5, tended to over-drive Wood Anderson seismometers. In response, low
amplification versions of this seismometer were deployed. However, it was expensive to
maintain a set of instruments that recorded only during the rare occurrence of $M_L > 5$.
An even more serious limitation was that the local magnitude scale was only defined for
California earthquakes that were recorded in California; it was not possible to compare
the size of earthquakes worldwide.

In 1945 Beno Gutenberg (Caltech) introduced the Surface Wave Magnitude M_S which is
defined as

$$M_S \equiv \mathrm{Log}\, A_{R20} - A(\Delta) \qquad\qquad 8.4$$

A_{R20} is the maximum amplitude (in microns) of the 20-s Rayleigh wave (typically on the
vertical component) recorded at distances exceeding 20º (about 2,200 km). Table 8.2
contains the values of A for a variety of distances

8-5

Δ (degrees)	−log A	Δ (degrees)	−log A
20	4.0	90	5.05
25	4.1	100	5.1
30	4.3	110	5.2
40	4.5	120	5.3
45	4.6	140	5.3
50	4.6	160	5.35
60	4.8	170	5.3
70	4.9	180	5.0
80	5.0		

Table 8.2. Distance correction factors for determination of M_S.

Rayleigh waves are often the largest wave group at these distances. Thus, A_{R20} is relatively easy to obtain using a photographic paper record, a meter scale and the response curve of standard seismographs (see Chapter 2). In principle, the measurement should be made after the record is narrow-band filtered at 20 s. Filtering these records would have been impractical in the 1940's; but fortunately 20 s Rayleigh waves are typically easy to identify on records (this arises from strong velocity dispersion in this period band). It's important to recognize that M_S is a measure of the Fourier amplitude spectrum at 20 s, while M_L is a maximum of the transient response of a 0.8-s sdof; that is these two scales measure different things. Nevertheless, Gutenberg and Richter adjusted A so that the two scales gave similar magnitudes for earthquakes where they could estimate both M_L and M_S. This meant that the scales converge at about $M_S \approx M_L \approx 6$. This can be seen in Figure 8.3, where M_L and M_S are plotted against each other. In particular, it appears that $M_L \sim \frac{2}{3} M_S$, or since these magnitude scales are logarithmic,

$$A_{Wood-Anderson} \sim A_{R20}^{\frac{2}{3}}$$. That is, **M_L and M_S are fundamentally different measures of an earthquake**.
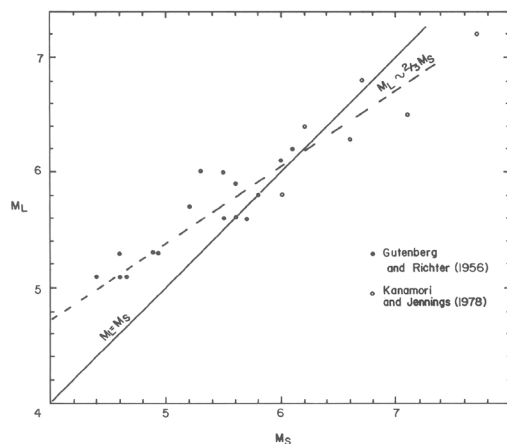

Fig. 8.3. Comparison of M_L and M_S for earthquakes reported for both scales. M_S is defined so that it gives comparable magnitude to M_L for earthquakes of about 6. From Heaton, T.H., Tajima. F., and Mori, A.W., 1986, Estimating ground motions using recorded accelerograms, Surveys in Geophysics, V. 8, pp 25-83.

Surface wave magnitude was a successful and commonly used scale for a large catalog of global earthquakes recorded in the 20th century (many older events were assigned magnitudes based on historic recordings. However, since the excitation of surface waves decreases exponentially with the depth of an event, M_S is not appropriate to characterize the size of earthquakes deeper than 20 km. To deal with this problem, Gutenberg, Richter, and Hugo Benioff (yet another famous Caltech seismologist) devised the **body wave magnitude scale, m_B**. This scale is based on teleseismic body waves (especially the P-wave between 30° and 90°). In this case, the maximum amplitude of the P-wave A_P and the duration of the pulse with the maximum amplitude T_P are measured.

$$m_B \equiv \log\left(\frac{A_P}{T_P A(\Delta, z)}\right)$$ 8.5

Where $A(\Delta, z)$ is a complex correction that depends on distance and source depth; it was calibrated to give comparable magnitudes to M_S for magnitudes near 7. Notice that the amplitude is normalized by the pulse duration. Richter says that this is a measurement of particle velocity which is used in the computation of kinetic energy.

The short-period body wave magnitude m_b is also commonly encountered in catalogs of earthquakes. This magnitude is also logarithmic and it depends on the maximum amplitude of teleseismic P-wave amplitudes as measured in the first 3 seconds of a short-period standard vertical seismometer (1-s natural period). m_b was designed to help discriminate the seismic signals from underground nuclear explosions from seismograms of natural earthquakes. That is, nuclear explosions happen very quickly and they have very impulsive P-waves, whereas earthquakes are typically more emergent; they take some time to get going; that is events with high m_b/M_S were marked as possible underground nuclear tests.

The development of magnitude scales has a long and complex history. There have been many other magnitude scales that have been devised. Unfortunately, they all suffer from the fact that they are unit less; they measure relative size, but they are not related to physical parameters of the earthquake. Hiroo Kanamori (yet another Caltech Seismologist) defined **energy magnitude M_W** to be

$$M_W \equiv \frac{\log W_0 - 4.8}{1.5}$$ 8.6

Where W_0 is the total radiated energy in joules (the W is for work). Kanamori defined M_W in this way so that it would be compatible with the energy-magnitude relation developed by Gutenberg and Richter (1956, BSSA, 32, 163-191). Unfortunately, estimation of radiated energy is a difficult problem, which makes the use of M_W cumbersome in practice. Kanamori noted that on average

$$W_0 \approx \frac{M_0}{2 \times 10^4} = \frac{\mu \bar{D} S}{2 \times 10^4} = \sigma_{eff} \bar{D} S = \sigma_{eff} P$$ 8.7

Where μ is the average crustal rigidity, about 40 GPa, which means that $\sigma_{eff} \approx 2 MPa$. σ_{eff} is the **effective stress** that converts slip per unit rupture area into radiated energy per unit rupture area. By substituting 8.7 into 8.6, Kanamori was able to approximate **M_W** as

$$M_W \approx \frac{\log M_0 - 9.1}{1.5}$$

8.8

$$\approx \frac{\log P + 1.5}{1.5} = \frac{2}{3}\log P + 1$$

Coincidentally, Thatcher and Hanks (1973, JGR, 78, 8547-8576) derived a general relation between M_0 and M_L for southern California earthquakes

$$M_L \approx \frac{\log M_0 - 9}{1.5}$$

8.9

Hanks noted that the M_L relation (eq.8.9) is very similar to Kanamori's approximate relation for M_W (8.8), and he convinced Kanamori to collaborate on a new definition of magnitude they called **moment magnitude $\mathcal{M}$** defined as

$$\mathcal{M} \equiv \frac{\log M_0 - 9.05}{1.5}$$

8.10

Where the 9.05 was obviously a political compromise (Hanks and Kanamori, 1979, JGR, 84, 2348-2350). The choice of the distinctive bold script M meant that it was awkward to find the appropriate symbol on word processors and hence many researchers began to call M_W moment magnitude. Unfortunately, some have mistakenly confused energy and moment (see Chapter 7). On a personal note, it drives me crazy that one of the most referenced papers in geophysics (Hanks and Kanamori, 1979) is just a simple 2-page paper to define moment magnitude, but most researchers who cite it use it incorrectly; M_W is defined to be Energy Magnitude. M_W **is NOT Moment Magnitude**.

The easy form for moment magnitude (technically it's Potency Magnitude) is that given in 8.8

$$\mathcal{M} \approx 1 + \frac{2}{3}\log S\overline{D}$$

8.11

Where the units are in meters. I have assumed that the average crustal rigidity is 35 GPa in deriving 8.11. This can cause some confusion, since some researchers use 40 GPa and others use 35 GPA. However, since the rigidity is not really a source parameter (see Chapter 7), this is not a significant issue. Since it makes the definition of magnitude simple, I suggest using 35 GPa.

As a convenient rough rule of thumb, a 30 km by 30 km rupture with a 1-m slip gives a potency of about 1 cubic km, which corresponds to a moment magnitude of about 7. From now on, I will refer to Potency Magnitude as simply magnitude. While it is, for all intents and purposes, the same as moment magnitude, potency is a much simpler concept and its use simplifies the math.

Figure 8.4 is a comparison of a variety of different magnitude scales (Heaton, Tajima, Mori, 1986, Surveys in Geophysics, 8, 25-83). It shows the average relationship between these scales and moment magnitude. This plot clearly shows that most magnitude **scales saturate** with increasing total energy (or potency). The scales based on shorter-period measurements saturate at smaller magnitudes and the scales based on long-period measurements saturate at larger magnitudes.
Unfortunately, magnitude saturation was poorly understood prior to the 1980's and there are important inconsistencies in reporting the sizes of older large earthquakes. For example, the 1906 San Francisco earthquake was assigned a surface wave magnitude, M_S, of 8.3 based on teleseismic recordings. In comparison, the 1964 Alaskan earthquake was

assigned an M_S of 8 ¼ . However, modeling of geodetic data from these earthquakes has indicated that the 1906 earthquake was $\mathcal{M} = 7$ ¾, while 1964 was $\mathcal{M} = 9$ ¼ ; that is the Alaskan earthquake was 1.5 units greater, or $10^{3/2 \cdot 1.5} = 180$ times larger. The use of the moment magnitude scale revolutionized the understanding of the sizes of the greatest earthquakes. You can find a table listing the largest historic earthquakes sorted according to their moment magnitudes at https://en.wikipedia.org/wiki/Lists_of_earthquakes.
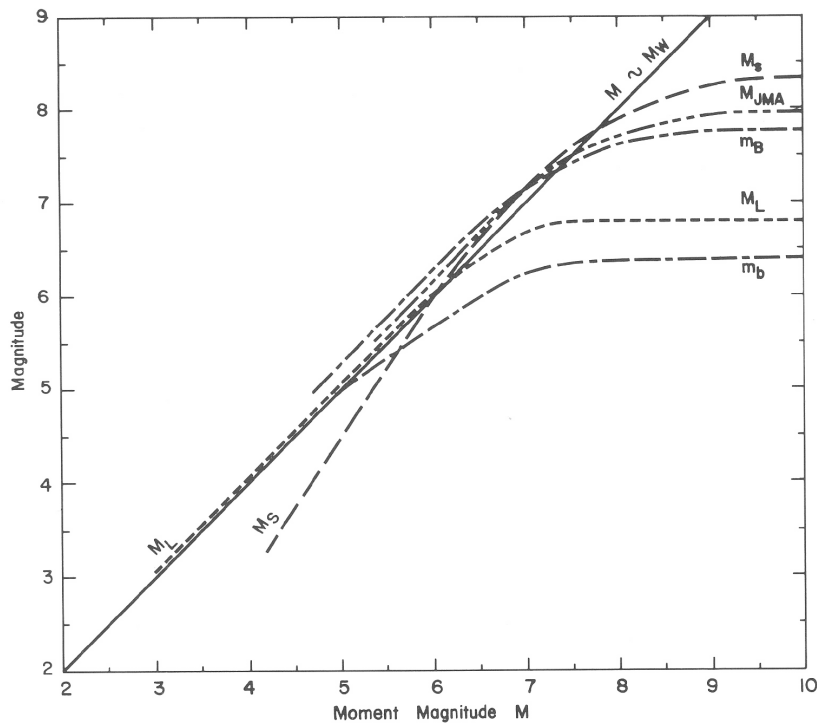

Figure 8.4. Comparison of a variety of magnitude scales (Heaton, Tajima, Mori, 1986, Surveys in Geophysics, 8, 25-83).

At this point, you should be asking yourself about this factor of $^2/_3$ that shows up in definition of moment magnitude; is there some physics behind this factor? One way to think of this problem is to consider a far-field P- or S-wave pulse of duration T and amplitude A. The potency (or moment) of the earthquake is then related to the area of the pulse, i.e., $P \sim AT$ (see Chapter 7). This implies that $A \sim \dfrac{P}{T}$. The assumption that most magnitude scales are logarithmic in amplitude, A, means that the definition of potency magnitude is consistent with other scales if $P \sim T^3$, in which case $A \sim P^{2/3}$, which implies that

$$\mathcal{M} \sim \log A \sim \tfrac{2}{3} \log P \qquad\qquad 8.12$$

As we will see in the following discussion, there are other ways to derive the $\tfrac{2}{3}$. Furthermore, this factor was in the relationship between magnitude and radiated energy that was derived by Gutenberg.

Gutenberg-Richter Relationship

Since the earliest days of compiling catalogs of earthquake magnitude, it has been noted that the number of earthquakes decreases with increasing magnitude and can be described as a power law known as the Gutenberg-Richter frequency-magnitude relationship (often shortened to just "Gutenberg-Richter").

$$\log(N) = A - bM \quad \text{or}$$
$$N = 10^A 10^{-bM}$$

8.13

This can be rewritten in natural logs as

$$\ln(N) = \log e \log(N) = A \log e - (b \log e) M$$
$$\text{or } N = \exp(A \log e) \exp\left[(b \log e) M \right]$$

8.14

Where N is the number of earthquakes larger than M, and A and b are constants that are typically called "the A value" and the "b value". This equation is the first form of Gutenberg-Richter and it is called the "cumulative form of Gutenberg-Richter" . A second form of this relationship describes the statistics of the number of events $N'\Delta M$ with magnitudes between M and $M + \Delta M$. In this case, $N' = \dfrac{dN}{dM}$. Now differentiating 8.13 with respect to M , $\dfrac{d(\log N)}{dM} = -b$. However, we also know that

$\dfrac{d(\log N)}{dM} = \dfrac{1}{N}\dfrac{dN}{dM} = \dfrac{N'}{N}$, so $N' = \dfrac{-bN}{\log e}$ from which it follows that

$\log N' = \log N - \log b - 0.362$ which provides the second form of the Gutenberg-Richter relation (called the interval form) that is written

$$\log(N') = a - bM$$

8.15

where $a = A + 0.362 + \log b + \log \Delta M$. Notice that the two forms, 8.13 and 8.15, only differ by a constant. If a large data set exactly fits the G-R relation, then the a- and b-values can be estimated using a visual inspection of a log-linear plot ($\log N$ vs M). However, it is more common to encounter catalogs that do not precisely fit this law, in which case it is easier to estimate the constant when plotting the first form. Most commonly, people use the first form. However, if you wish to visually judge whether a data set fits G-R, it is probably best to use the second form; it reveals the details better.

Interpreting the meaning of the G-R relation has been the subject of a huge number of papers, but to this day, there still seems to be a wide variety of speculative explanations. The a-value is a constant that gives the overall rate of activity. The b-value describes the relative frequency of large and small events. Later in this chapter, I argue that the b-value arises from the chaotic dynamics that are caused by the low dissipation energy in earthquakes.

It is important to understand that, given a suite of earthquakes, the estimates of the a- and b- values is dependent on the magnitude scale that was used to parameterize the events. For example, I argued that $M_L \sim \frac{2}{3} M_S$. This means that using M_L to determine b-value will give a different estimate than if the catalog uses M_S . Fortunately, most networks that prepare earthquake catalogs attempt to enforce the use of moment magnitude. Unfortunately, many catalogs of historic earthquakes contain a variety of different magnitude estimates.

The G-R relation is also referred to as a *Pareto Distribution*. This name is common in economics, where it has long been observed that personal wealth is described by a Pareto distribution; a few people have most of the overall wealth in society. Pareto distributions often arise in self-organizing systems. I have found a very nice overview lesson on power laws by Mark Newman (U. of Mich.) which can be found at http://www-personal.umich.edu/~mejn/courses/2006/cmplxsys899/powerlaws.pdf

We can use the Gutenberg Richter relation to investigate the following common question. Which magnitude earthquakes are responsible for most of the motion of the Earth's tectonic plates? Is it the infrequent large events, or the far more numerous smaller earthquakes? To estimate the total inelastic strain integrated over the volume, total Potency or P_{total} , resulting from suite of earthquakes between magnitudes M_{min} and M_{max} , we perform the following integral

$$P_{total} = \int_{M_{min}}^{M_{max}} N'P(M)dM \qquad\qquad 8.16$$

Where $P(M) = 10^{\frac{3}{2}(M-1)}$ (see 8.8) and N' is given by 8.15. Performing the integral, we obtain

$$P_{total} = 10^a \int_{M_{min}}^{M_{max}} 10^{(1.5-b)(M-1)}dM$$

$$= \frac{10^a}{(1.5-b)\ln 10}\left[10^{(1.5-b)M} \right]_{M_{min}}^{M_{max}} \qquad\qquad 8.17$$

If $M_{min} = -\infty$, then

$$P_{total} = \frac{10^{a+(1.5-b)M_{max}}}{(1.5-b)\ln 10} \sim C10^{(1.5-b)M_{max}} \qquad\qquad 8.18$$

Analysis of the global catalog of earthquakes provides estimates that $b \approx 1$, so $P_{total} \sim C10^{\frac{1}{2}M_{max}}$. This means that the total slip increases by a factor of 3.16 for every 1-unit increase in M_{max} ; clearly **the large earthquakes account for the majority of the deformation. This fact means that it is difficult to use historic seismicity catalogs to estimate tectonic deformation rates. The largest events dominate the total deformation estimate, but they are so infrequent that there are too few events to do statistics. Furthermore, Gutenberg Richter converges to a finite tectonic rate only in the event that the power law is truncated at the largest events. If there is no largest size event, then 8.18 is divergent. Although attempts have been made to estimate the maximum event, there is currently no accepted methodology to**

truncate the Gutenberg-Richter Relation. As you read further in this chapter, I think that you will see that estimation of a largest event is not a solvable problem.

Notice that if $b > 1.5$, then the situation reverses and more of the deformation is caused by the small events. Taken literally, $b > 1.5$ means that there is so much deformation from the smallest events that there does not need to be any larger events, that is, the plate boundary is creeping.

One notable aspect of the G-R relation is the way it scales rupture area. In particular, it has been observed that seismic moment scales with rupture area as

$$M_0 \approx 1.2 x 10^{15} S^{3/2} \hspace{3cm} 8.19$$

Where S is rupture area in km^2, and moment is in N-m see Figure 8.5 from Kanamori (1980). Writing this in terms of potency,

$$P \approx 3 x 10^4 S^{3/2} \hspace{3cm} 8.20$$

Where S is still in km^2. Converting S to m^2, we obtain.

$$P \approx 3 x 10^{-5} S^{3/2} \hspace{3cm} 8.21$$

This allows us to infer the total rupture area of an ensemble of earthquakes as a function of M. Combining 8.21 and 8.11, we obtain

$$M \approx 1 + \frac{2}{3} \log P \approx 1 + \frac{2}{3} \log \left(3 \times 10^{-5} \right) S^{3/2}$$
$$= \log S - 2.02 \hspace{3cm} 8.22$$

Or

$$S \approx 10^M 10^{2.} \hspace{3cm} 8.23$$

This means that the sum of the rupture areas of earthquakes with magnitudes between M and M+ΔM is

$$S' = N' \Delta M S(M) = 10^{a-bM} 10^{M+2.} = 10^{a+2.} 10^{(1-b)M} \hspace{2cm} 8.24$$

This means that S' is a constant if $b = 1$. That is, **the total rupture area is the same for all earthquakes in a magnitude interval** ΔM. That is, if the b-value is 1, then the total rupture area is the same for integrated area of each magnitude. That is the sum of the rupture area of all 2's is the same as the rupture area of all 3's, is the same as all 4's, etc. What this means is that given a b-value of 1, and given that a point has just experienced slip, then it is equally likely that it could have come from any magnitude earthquake. Given that different magnitude earthquakes have different slips, any slip is as likely as any other. This is only true in a logarithmic sense. That is, a fault is equally likely to experience slip between D and C x D, regardless of the value of D and a constant, C (e.g. given, slips between 1 and 2 mm are just as common as slips between 4 and 8 m. For example, if you sum the rupture areas of 10 M 6 earthquakes, they would equal the rupture area of a M 7. This allows us to answer the following question. "Given that a point on a fault experiences a rupture, what is the most likely magnitude of earthquake that caused the rupture?" Apparently, if $b = 1$, then all magnitude earthquakes are

equally likely to have been the cause of the rupture. If $b < 1.0$, then the given rupture is most likely caused by a large event, whereas if $b > 1.0$, then it's most likely that the event was caused by a small event.

From the previous discussion, we can conclude that if there is a point on a fault that fails in an earthquake, then it is equally likely that the causative event was any magnitude. Later I show that the average slip in an event increases as $10^{\frac{1}{2}M}$ (see 8.78), which means that, "**Given that a point on a fault experiences a rupture, then all values of slip at that point are equally likely**."

Stress Drop and Strain Change

The term, stress drop, is often used in seismology and unfortunately, it has been defined in many ways (often confusing). In the following discussion, **I will define stress drop to be the change in shear traction on a fault that is the result of shear displacements that occur in an earthquake**. That is,

$$\Delta\sigma(x,y) \equiv \sigma_0(x,y) - \sigma_1(x,y) \qquad 8.25$$

Where σ_0 is the initial stress (often called the prestress) and σ_1 is the final stress. This definition of stress drop is often referred to as the **static stress drop**. As far as I am concerned, this is the only unambiguous definition for stress drop, although it seems that many (perhaps most) of my colleagues use other definitions.

The easiest way to think of static stress drop at a point on a fault is that it is the change in elastic strain time the rigidity, or $\Delta\sigma \equiv \sigma_0 - \sigma_1 = \mu(\Delta\varepsilon)$, where $\Delta\varepsilon$ is the amplitude of the change in shear strain in the volume immediately adjacent to the rupture surface. Unfortunately, things are not quite so simple since $\Delta\sigma$ and $\Delta\varepsilon$ are actually tensors that vary over space such that they satisfy the equations of static equilibrium ($\sigma_{ij,j} = 0$). However, in these class notes, I am not very interested in the details and I will use appropriately simple definitions. In particular, I **assume that the average stress drop of spatially complex ruptures is approximated by rectangular faults with homogeneous stress drop**. The adequacy of this approximation is discussed by Das (1988, Relation between average slip and average stress drop for rectangular faults with multiple asperities, BSSA, 78, 924-930) and Noda et. al. (2013).

Notice that all stress and strain changes scale linearly with slip. That is, whatever average stress drop corresponds to a particular distribution of slip, the average stress drop doubles when we double all of the slips.

I will use the relations between average slip, $\overline{D}$, rupture length, L, and rupture width, W , for a rectangular fault in a homogeneous half space as reported by Parsons (Parsons, Ian Dennis, 1988, *The Application of the Multigrid Method to the Finite Element Solution of Solid Mechanics Problems.* Dissertation (Ph.D.), California Institute of Technology. doi:10.7907/MVMM-ED69. https://resolver.caltech.edu/CaltechETD:etd-08102006-090020et. al. This is actually a difficult solid mechanics problem and there are several flawed solutions that are in use. I believe that the Parsons solution is reliable.

In these discussions I will assume that the rupture surface can be approximated by a rectangle with a long dimension of L and a short dimension of W. In many cases, L becomes the along-strike dimension, whereas W is the down-dip dimension. Further, I will assume that

$$\overline{\Delta\sigma} \approx \mu\left(C\frac{\overline{D}}{W}\right)$$

$$= \mu\left(C\frac{P}{LW^2}\right)$$

8.26

Or alternatively,

$$\overline{\Delta\varepsilon} \approx \left(C\frac{\overline{D}}{W}\right)$$

$$= \left(C\frac{P}{LW^2}\right)$$

8.27

Where $P = LW\overline{D}$ is the Potency (the volume integral of inelastic shear strain in an earthquake, see Chapter 7), C is a dimensionless constant that depends on the aspect ratio of the rupture, $\frac{L}{W}$, the depth of burial of the rupture z_{top}, and the rupture dip and rake angles. Of course, since I am assuming a homogeneous half-space, it is a trivial change to write 8.26 in terms of seismic moment $M_0 = \mu P$. Parsons (1988) shows that for a fault with rupture that reaches the Earth's surface,

$$C \approx 1.39\left[2.3^{\left(1-\frac{L}{W}\right)}\right] + 0.65$$

8.28

And for a deeply buried fault

$$C \approx 1.27\left[2.3^{\left(1-\frac{L}{W}\right)}\right] + 1.28$$

8.29

For ruptures with shallow burial depth, C changes quickly as a function of $\left(\frac{z_{top}}{W}\right)$; C is within 95% of its value for deeply buried when $\left(\frac{z_{top}}{W}\right) = 0.16$. This rapid change in C with burial depth is caused by the fact that, in an elastic model, there are very large strains and stresses in the region between the top of the fault and the free surface. In reality, it is highly questionable whether such large stresses develop in the shallow region just above a rupture that does not quite reach the surface. For simplicity, I classify events as either "surface rupturing" or "deeply buried"; that is, I assume that C is described by 8.29 for any earthquake that did not have surface rupture. I assume a uniform rigidity of $\mu \approx 35GPa$, which is taken to represent average properties in the upper Crust.

Notice that 8.21 can be rewritten as

$$\overline{D} \approx 3\times10^{-5}\sqrt{S} = 3\times10^{-5}\sqrt{LW}$$

8.30

If I assume that ruptures are geometrically similar, then $\frac{L}{W} \sim a$, where a is the aspect ratio between rupture length and rupture width. In this case,

$$\bar{D} \approx 3\times10^{-5}W\sqrt{a} \qquad\qquad 8.31$$

If typical ruptures are as long as they are wide, then $\bar{D}/W \approx 3\times10^{-5}$. **That is, the typical slip to length ratio of earthquakes is about 3 parts in 100,000.** Substituting this into 8.26 gives

$$\Delta\sigma \approx \mu C\times3\times10^{-5} \qquad\qquad 8.32$$

And

$$\overline{\Delta\varepsilon} \approx 7.6\times10^{-5} \qquad\qquad 8.33$$

If the rupture is deeply buried and if the aspect ratio is about 1, then $C \approx 2.55$. If $\mu \approx 35GPa$, then we conclude that $\Delta\sigma \approx 2.7MPa$ and $\overline{\Delta\varepsilon} \approx C\times3\times10^{-5}$. While most current analysis uses stress drop as a parameter, I find that it's easier to use strain drop since 1) it's directly related to the observables (Potency and Rupture area), and 2) it is a unitless quantity.

If you inspect Fig. 8.5, you will see diagonal lines with labels of different stress drops given in bars. A bar is defined to be mean atmospheric pressure at sea level, and it is approximately equal to 100 kPa. The use of bars to describe stress was common decades ago, but almost all research journals now require the use of Standard International (SI) units, where stress has units of Pascals, Pa=N/m^2.
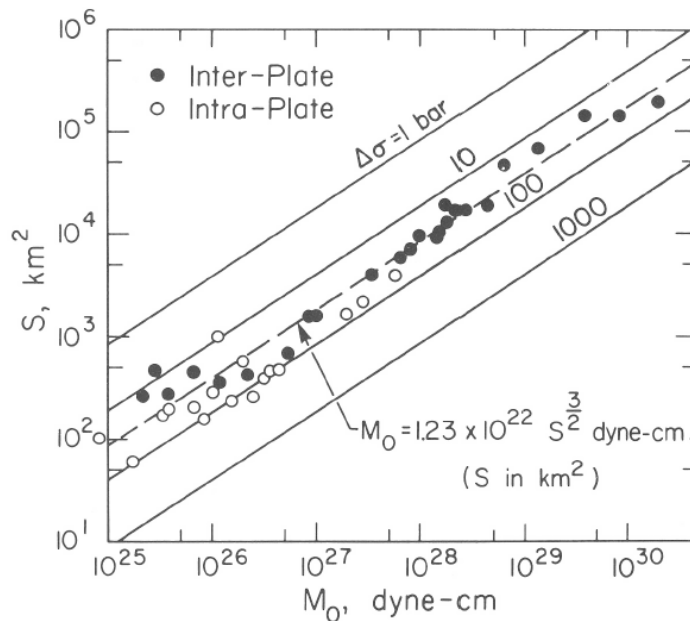
Fig. 12.   Relationship between fault area and seismic moment. The solid lines denote constant stress drops and the dashed line ($M_0 = 1.23 \times 10^{22}$ S$^{3/2}$ dyne-cm; S in km^2) signifies a stress drop of about 30 bars. Solid points denote interplate earthquakes and open circle denote intraplate earthquakes. This figure is taken from Kanamori (1980) and is used with permission from the author.

Figure 8.5. This is a copy of a Figure in Heaton, Tajima, and Mori,1986, Surveys in Geophysics, V. 8, pp 25-83, which is a copy of a very famous figure by Kanamori, 1980.

Earthquakes modeled as Brittle Cracks

In the discussion in the previous section about stress drop, we approximated an earthquake as having a constant change in stress on a planar area. This problem is almost identical to the problem of introducing a frictionless crack into a uniformly stressed elastic material. This allows us to borrow the extensive research developed to describe the failure of brittle materials. Crack-like models of dynamic rupture were common in the 1960's and 70's (e.g., Kostrov, 1966, Self-similar problems of propagation of shear cracks, J. App. Math. Mech., 28, 1077-1087; Madariaga, 1976,).

In Chapter 3, I described the solution of an instantaneous pressure change in a spherical cavity (3.82). In that solution, the cavity expanded when the pressure increased. In addition, the expansion looked like the solution to an impulse of momentum on a 58% damped sdof. That is the cavity oscillated about its new static equilibrium and it radiated wave energy as it oscillated. In the case of a frictionless crack with a step change in shear traction, the crack also oscillates about its new equilibrium. I don't know any closed form solution to this problem, but Brad Aagaard made some finite element simulations of a frictionless crack and he found that it had radiation damping of about 20% (unpublished research). Fault friction differs from the frictionless case in that once the slip drops to zero, the friction returns and presumably the motion stops. Thirty years ago there was discussion about whether the slip would stop at its dynamic maximum, or whether it would stop at its static equilibrium. This was referred to as the "overshoot" problem. In the past few decades, researchers construct numerical simulations of faults with very specific friction laws that are more complex than simple instantaneous change in traction.

True analytic solutions of spontaneously propagating ruptures are rare. One simple approach to the problem of a steadily propagating semi-infinite crack (shown in Figure 8.6) is to start with the spatial solution for a static semi-infinite crack. To obtain the solution for a propagating crack, merely move the observer at a steady rate opposite to the direction of the crack propagation; the transforming velocity is scaled by the shear-wave velocity (similar to a relativistic Lorentz transformation).

The static solution for a semi-infinite shear crack with a constant stress drop is

$$\sigma_{xz} \sim \frac{K}{\sqrt{r}} f_{xz}(\varphi) \qquad\qquad 8.34$$

Where K is the **stress intensity factor** (a measure of the sharpness of the crack tip), r is distance from the crack tip, and $f_{xz}(\varphi)$ is a scalar function to account for the polar coordinate φ, which is the angular position of the observer as measured with respect to the rupture plane. This solution assumes that everything is uniform in the y direction. That is the crack extends infinitely in y direction. For a mode III crack, $K_{III} = \sqrt{2\mu G}$, where G is the *fracture energy*, which is the energy per unit length required to extend a semi-infinite crack. If the fracture energy becomes small, then the tip of the crack approaches the linearly elastic solution, which has infinite stress at the crack tip. That is, if the crack tip

is linearly elastic, then even small applied stresses will cause very large stresses at the crack tip.

The work expended on the crack of length L by slippage D is just $DL\left(\sigma_f + \frac{1}{2}\Delta\sigma\right)$. In a true crack, the sliding surface is frictionless $\left(\sigma_f = 0\right)$ and the total change in potential energy is

$$\Delta E_{potential} \approx DL\left(\frac{1}{2}\Delta\sigma\right) \approx \left(\frac{C}{2\mu}\right)L\Delta\sigma L\Delta\sigma = \left(\frac{C}{2\mu}\right)L^2\Delta\sigma^2 \quad 8.35$$

When a mode III crack is extending in a steady-state way (constant rupture velocity), then the rate change of potential energy is balanced by the rate of change of fracture energy, or $\dot{E}_{fracture} = \dot{E}_{potential}$. The rate of change of fracture energy is just

$$\dot{E}_{fracture} = GV_R \quad\quad 8.36$$

where V_R is the rupture velocity. The rate of change of potential energy is calculate by differentiating 8.35 with respect to time, or

$$\Delta\dot{E}_{potential} \approx \left(\frac{C}{\mu}\right)L\dot{L}\Delta\sigma^2 = \left(\frac{C}{\mu}\right)LV_R\Delta\sigma^2 \quad\quad 8.37$$

which means that

$$G = \left(\frac{C}{\mu}\right)L\Delta\sigma^2 \quad\quad 8.38$$

For an equilibrium crack, the shear stress on the fault plane scales as

$$\sigma_{xz} \sim \frac{L\Delta\sigma}{\sqrt{x}} \quad\quad 8.39$$

For a propagating crack, this transforms to

$$\sigma_{xz} \sim \frac{L\Delta\sigma}{\sqrt{x'}} \quad\quad 8.40$$

Where x' is time-transformed distance given by

$$x' = \frac{1}{\sqrt{1-\left(\frac{V_R}{\beta}\right)^2}}\left(x - V_R t\right) \quad\quad 8.41$$

As the rupture velocity approaches the intrinsic shear wave velocity, the solution appears to be compressed in space. Stress at the equilibrium crack tip is singular. When the crack runs at the shear-wave velocity, it becomes very singular.

In crack theory, an expanding crack under uniform load tends to continue to expand indefinitely. That is, once it becomes unstable, the only way to stop it is

with high fracture energy. The longer the crack length, the harder it is to stop, since the required fracture energy increases as $L(\Delta\sigma)^2$.

The laboratory-measured values of fracture energy of materials depend strongly on the brittleness of the material. Although mono-crystalline diamond is the hardest known material, it is also extremely brittle and only 5.5 J/m^2 is required to propagate a mode I fracture along a cleavage plane. This means that diamond is very strong at the microscale, but you wouldn't want to construct anything large out of a single diamond crystal; if a crack was initiated, then it would not take much average stress to propagate a crack through the rest of the material. Glass is another brittle material and its measured fracture energies range from 3.5 to 5.5 J/m^2.
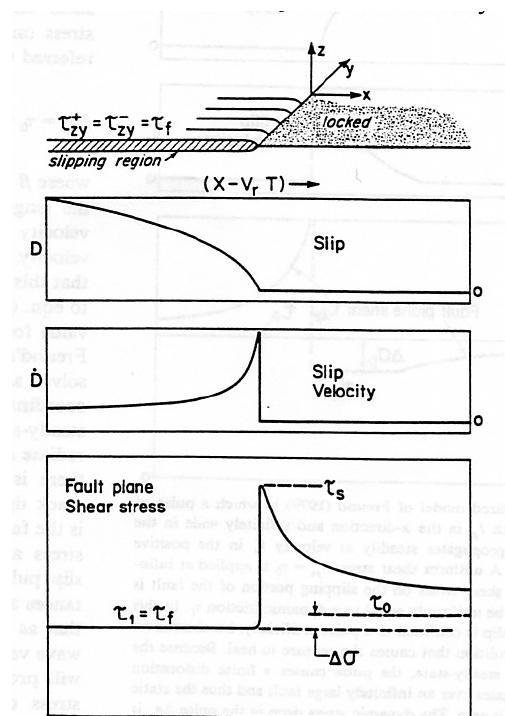

Figure 8.6. Simplified view of an earthquake as an expanding crack. This is often called a semi-infinite crack since it is uniform in the y coordinate.

Polycrystalline materials tend to have higher fracture energies because the fracture surfaces are complex and more molecular bonds must be severed than for a similar dimension of single crystalline material. Concrete has a fracture energy of about 190 J/m^2, and most polycrystalline rocks (e.g., granite) have fracture energies between 100 and 1,000 J/m^2.

In contrast, earthquake energies are much, much larger than measured values of G. For example, 8.7 gives estimates of radiated energies of

$$\frac{E_R}{S} \approx \sigma_{eff} D \approx 2D\,\mathrm{MJm}^{-2} \qquad\qquad 8.42$$

8-18

That is, the radiated energy for a typical 1-m slip is about 2 MJ/m^2, which is at least 1,000 times larger than measured fracture energies measured in a laboratory. Or, alternatively, a 10 km by 10 km rupture with a stress drop of 3 MPa has a fracture energy of 30 MJ/m^2 (see 8.38).

This is a **MAJOR PROBLEM** with interpreting earthquakes using simple constant-stress-drop cracks. That is, **constant stress drop cracks are NOT SCALE FREE since they imply that fracture energies (the work to stop ruptures) increases as** L . The notion that the earthquake dynamics are controlled by fracture energy processes at the crack tip that can be 10's of km away seems totally implausible to me. Unfortunately, this idea is central to many popular earthquake physics studies as I will discuss shortly.

Earthquake Similarity

Notice that Fig. 8.5 shows that the average stress drop of 2.7 MPa is a kind of universal relation that applies over a very wide range of earthquake sizes. In essence, the observation is that average slip correlates with rupture dimension as stated by 8.31. The apparent universality of this relationship was the motivation for a class of earthquake scaling relations that are often referred to as "self-similar rupture models." The use of the term, self-similar, has created significant confusion, and I recommend using it only when you mean it.

Self-similar is usually used in physics and mathematics to describe **objects that look similar, independent of the magnification** at which the object is observed. In particular, **fractal objects that have similar complexity at all scales are referred to as self-similar.** That is, they look like themselves independent of the magnification. In contrast, two objects are **called similar when they look identical after an appropriate scale change**. For example, squares are all geometrically similar; all squares are identical if the length of a side is normalized to unity (remember similar triangles that you learned in high school).

Another related term that you might encounter is **self affine**. Objects that are mapped in multiple dimensions are typically called self-affine when linear scale changes in some of the dimensions result in objects that are self-similar. Again, this term is common in the study of fractals.

Aki (1967, Scaling law of seismic spectrum, JGR, 72, 1217-1231) proposed that all earthquakes were similar; he did not call then self-similar. That is, he suggested that different ruptures could be seen to be identical except for the length of the rupture. Aki's argument was simple and compelling. He argued that all earthquakes were fundamentally controlled by the same 3 MPa sudden drop in shear stress on the fault surface. That is, he argued that the stress drop is a fundamental (and universal) description of the friction on the fault. The initial stress σ_0 is assumed to be the strength of the fault, and the sliding friction, σ_f , is assumed to be equal to the final stress σ_1. Later this general model has been called "self-similar" [sic] constant-stress-drop model of earthquake ruptures. It has been the basis of numerous models to explain seismic data (especially the shape of Fourier amplitude spectra). According to Aki's conjecture, the primary difference between different earthquakes is the area and aspect ratio of a rupture. That is, once the rupture

length is chosen, then all other dynamic parameters can be inferred. That is, all of the dynamics are controlled by physics that is independent of the scale length of the earthquake. In this way these models are **scale invariant** (not really, though, because fracture energy increases with rupture dimension).

In the following pages I will describe the standard "self-similar" [sic] source model so that you will understand the origin of numerous terms that are commonly used. However, I warn you now that I view that these models are seriously flawed, and their use typically decreases real understanding. Furthermore, I will develop arguments that show that earthquakes are not scale invariant.

As we saw in Chapter 7, the displacements $u(t)$ from far-field S-waves radiated by a sliding fault is proportional to the potency rate, $\dot{P}$. Now so long as $P(t)$ is monotonically increasing with time, then $\dot{P}(t)$ *is a strictly positive function*. Now we can conclude that

$$P = \int_0^T \dot{P}\,dt \sim \int_0^T u\,dt \qquad\qquad 8.43$$

Where T is the total duration of the earthquake. While we could evaluate P by a simple integration of $u(t)$ in time, it has become customary to perform this integration in the Fourier transformed domain. In particular,

$$\lim_{\omega \to 0}\left|\tilde{u}(\omega)\right| = \left|\int_{-\infty}^{\infty} u(t)\,dt\right| \sim P \qquad\qquad 8.44$$

That is, the earthquake potency is proportional to the zero-frequency-amplitude of the Fourier transform of the far-field S-wave. Of course, one must correct for radiation pattern and path effects (e.g., transmission and reflection coefficients and geometric spreading). In fact, the seismograms of what we usually call the S-wave are actually comprised of a combination of many different rays that sum together to comprise the "S-wave group." In order to deduce potency, P, from this data, one should really account for all of these different rays. However, it's far simpler to just take a Fourier transform of the S-wave group and to then find the amplitude at zero frequency. It's usually assumed that the complexities from many rays serves to make the phase spectrum look random while there is little systematic effect on the overall amplitude spectrum. While this assumption is pervasive, I am unaware of any study that provides convincing evidence to support it. For the purposes of the following discussion, though, I will assume that the spectra of real earthquakes are the same as the spectra of far-field waves radiated from a shear fault located in a homogeneous, isotropic whole space (***caveat emptor***).

Suppose that $u(t)$ is a strictly positive function of time and having a duration of T_c, then

$$\left|\tilde{u}(0 \le f \le f_c)\right| = \int_0^{T_c} u(t)\,dt \sim P \text{ and } \left|\tilde{u}(f > f_c)\right| < \left|\tilde{u}(f_c)\right| ,$$

where $f_c = \dfrac{1}{T_c}$ is the "corner frequency." As an example, consider a far-field S-wave that consists of a simple rectangle, or

$$u(t) = \Pi(t) \equiv \begin{bmatrix} 0 & t < -\dfrac{1}{2} \\ 1 & -\dfrac{1}{2} < t < \dfrac{1}{2} \\ 0 & t > \dfrac{1}{2} \end{bmatrix} \qquad 8.45$$

(see chapter 1 for discussion of the rectangle function). Now $\tilde{u}(f) = \mathrm{sinc}(f) \equiv \dfrac{\sin f}{f}$.

$\tilde{u}(f = 0) = 1$, which is the integral of the rectangle function. Furthermore, $\left| \tilde{u}(f \gg 1) \right| \to \dfrac{1}{f}$. The high-frequency asymptote, $\left| \tilde{u} \right| = \dfrac{1}{f}$ intersects the low-frequency asymptote, $\left| \tilde{u} \right| = 1$, at the corner frequency, $f_c = 1$, which is the reciprocal of the duration of the rectangle function. While f_c is well known for a rectangle function, it is something that must be determined when it is used to characterize an S-wave group. It is customary to plot the amplitude spectrum of the wave group on a log-log graph and to fit straight lines to the low frequencies and the high frequencies. The intersection of the two lines provides a measurement of the corner frequency.

We can use the corner frequency measurement to estimate the rupture dimension,

$$L \approx \frac{1}{2} T_c V_R = \frac{1}{2} \frac{V_R}{f_c} \qquad 8.46$$

where I have assumed that the earthquake duration is approximately twice the time it takes for a rupture front to sweep across the rupture surface. The factor of two is to account for the average duration of the slip. Assuming that $L \approx W$, $C \approx 2.55$, and applying 8.26 we obtain

$$\overline{\Delta\sigma} \approx 20\mu P \left(\frac{f_c}{V_R} \right)^3 \qquad 8.47$$

Or alternatively,

$$\overline{\Delta\varepsilon} \approx 20 P \left(\frac{f_c}{V_R} \right)^3 \qquad 8.48$$

Notice that we made a questionable assumption that the rupture velocity is a constant, independent of the earthquake size[1]. Furthermore, the stress drop is very sensitive to our funky measurement of the duration of the S-wave, f_c^{-1}, since the stress drop depends on the cube of the corner frequency; if your measurement of corner frequency is off by a factor of two, then your stress drop estimate is off by a factor of eight. Although I don't recommend it, 8.47 can be used to estimate average stress drop (or strain change) by simply measuring the asymptotic behavior of Fourier amplitude spectra. If you are genuinely interested in stress drop, then I strongly recommend that you obtain it from the types of finite-fault slip models that were described in Chapter 7. In particular, finite-

---

[1] This is an important assumption of the Brune source model, but the evidence is that the duration of slip at any point is small compared to the time it takes a rupture to propagate on a rupture surface.

element models can be used to calculate the spatial distribution of stress drop (or strain change) based on the spatial distribution of slip.

Aki's conjecture of universal similarity (1967) suggested that all earthquakes were identical dynamic processes except for the rupture dimension. That is, he suggested that all events had the same stress drop and that everything else could be deduced from this stress drop and the rupture dimension. In terms of the Fourier amplitude spectrum, this means that everything is controlled by the corner frequency, which is a description of the rupture dimension. Now if there is only one variable in our dynamics problem (the rupture length), then the shape of the Fourier amplitude spectrum must be described by only one variable, f_c . In this case, length is mapped to corner frequency through the rupture velocity. Aki argued that the spectrum must be of the following form

$$|\tilde{u}| \sim \frac{1}{\beta r} \frac{P}{1 + \left(\dfrac{f}{f_c}\right)^{\alpha}} \qquad\qquad 8.49$$

where α is a constant that is determined by the dynamics of the rupture process. Note that Aki used M_0 instead of P, so his equations look a little different. The r^{-1} is to account for geometric spreading of a far-field body wave, and the β^{-1} is an impedance factor that tells the size of radiated waves relative to the near-source term (look back at the far-field terms in 7.45). We can now estimate the energy in the radiated wavefield as follows. The power $\dot{E}_R$ flowing through an increment of area ΔS in the radiated S-wave is (see Chapter 3, eqn 3.55)

$$\frac{\dot{E}_R}{\Delta S} = \sigma \dot{u} = \rho \beta \dot{u}^2 \qquad\qquad 8.50$$

Since this energy is traveling radially to great distance, We can calculate the radiated energy traveling through an element of surface area dS by integrating 8.50 with respect to time. That is,

$$\frac{E_R}{\Delta S} = \rho\beta \int_0^\infty \dot{u}^2 dt = \rho\beta \int_0^\infty \tilde{\dot{u}}^2 df = \rho\beta (2\pi)^2 \int_0^\infty (f\tilde{u})^2 df \qquad 8.51$$

Our integral over time is transformed to an integral over frequency using Parseval's theorem. The integral over frequency can be broken into two separate integrals, the first over the constant amplitude at frequencies less than f_c , and the second over the frequencies higher than f_c . Combining 8.49 and 8.51, we obtain

$$\left|\frac{E_R}{\Delta S}\right| = \frac{\rho}{r^2 \beta} (2\pi)^2 P^2 \int_0^\infty \frac{f^2}{\left[1 + \left(\dfrac{f}{f_c}\right)^{\alpha}\right]^2} df \qquad\qquad 8.52$$

Unfortunately, the solution to this integral is a hypergeometric function, which is not very useful for this discussion. Fortunately, the integral can be solved in closed form if $\alpha = 2$ (the commonly used f^{-2} model). In this case

$$\frac{E_R}{\Delta S} = \frac{\pi^3 \rho}{\beta r^2} P^2 f_c^3 \qquad 8.53$$

provided that $\alpha \geq {}^3\!/_2$; the energy at high-frequencies is infinite for any $\alpha < {}^1\!/_2$. Notice that 8.47 can be used to conclude that

$$f_c \approx \left(\frac{V_R^3}{20\mu} \frac{\overline{\Delta\sigma}}{P} \right)^{1/3} \qquad 8.54$$

Or even more conveniently,

$$P f_c^3 \approx \frac{V_R^3}{20} \overline{\Delta\varepsilon} \qquad 8.55$$

or

$$\overline{\Delta\varepsilon} \approx \frac{20 P f_c^3}{V_R^3} \qquad 8.56$$

so 8.53 can be written

$$E_R \approx \pi^3 \rho V_R \left(\frac{V_R}{\beta} \right)^2 \Delta\sigma P \qquad 8.57$$

While the issue of spectral scaling for $f < f_c$ is well studied and non-controversial, the spectral characteristics for $f > f_c$ have been the focus of a huge number of studies, and unfortunately there is a depressing lack of consensus about spectral scaling for radiated high frequencies. Aki (1967) presented both an $\alpha = 2$ model and also an $\alpha = 3$ model. He argued that the apparent geometric similarity of ruptures (approximately scale independent stress drop) was a sign that models should be appropriately simple, which he interpreted to mean that spectral scaling should have power-law scaling with a simple integer in the power law. He said he did not know whether f^{-2} or f^{-3} was more appropriate (f^{-1} leads to infinite energy), but based on sparse data, he had a preference for f^{-2} (often called the omega squared model) . Figure 8.7 shows Aki's two hypothesized spectral models.

While Aki's hypothesized model was scale independent in many ways, it is clearly not scale independent when it comes to fracture energy. That is, Aki's model is essentially a constant stress drop crack model. This is the class of models that require a fracture energy that increases with rupture dimension. Later I will show that slip-pulse models can be constructed in a way that they produce scale-independent stress drop and also scale-independent fracture energy.
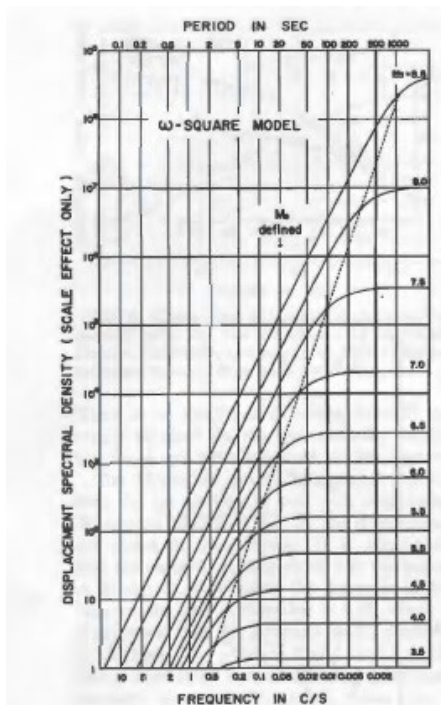
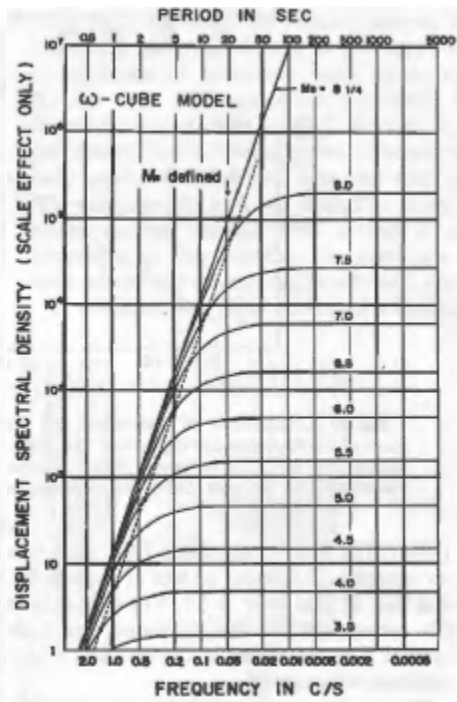Fig. 3. Dependence of amplitude spectral density of earthquake magnitude M_s for the ω-square model.

Fig. 4. Dependence of amplitude spectral density on earthquake magnitude M_s for the ω-cube model.

Figure 8.7 Aki's spectral models. The f^{-2} model on the left is often referred to as a self-similar (sic.) model. (Aki, K., 1967, J Geophys. Res. Scaling law of Seismic Spectrum, v. 72, 1217-1231)

Brune's Spectral Scaling Model

Brune's 1970 paper, Tectonic stress and the spectra of seismic shear waves from earthquakes (J. Geophys. Res., v.75, 4997-5009) is one of the most cited papers in all of earthquake research (approximately 5,000 citations according to Google Scholar). This paper is basically an extension of Aki's 1967 paper on universal scaling. In this paper, Brune argues that the dynamics of earthquake slips are controlled by an effective shear stress, σ_e, that accelerates the faces of an earthquake fault. That is, he argued that the fault appears as if it's a free surface with an instantaneously applied shear traction that is equal to σ_e. Brune then introduced the slip history for a fault that experiences an instantaneous step in shear traction. He uses this solution to infer the spectral characteristics of radiated far-field waves. That is, Brune argues that he introduced dynamics into the types of spectral models developed by Aki. Brune argues that one can derive the effective stress by appropriate modeling of the envelope of the Fourier amplitude spectra of seismograms. Notice that in the previous discussions, I was careful to describe the solutions using either change in strain, or in terms of stress drop. If the material is linearly elastic, strain and stress are linked by the shear modulus. The seismograms used to construct spectra only measure lengths and time, so the parameter strain change is natural. In contrast, Brune's analysis is formulated in terms of stress drop. In Brune's model, the stress drop is defines the shear tractions that accelerate the sides of the fault.

I will now try to describe Brune's 1970 spectral model, although that is a difficult task. In particular, what is currently called the Brune model is different from what is presented in his 1970 paper. In addition, Brune's induction about rupture physics has some rather glaring errors that are seldomly acknowledged.

I will begin by saying that the spectral envelopes of a wide range of seismograms can be approximately described with the Brune model; this is what makes this model so popular. Unfortunately, the glaring physics errors also make this paper confusing; there have been countless misuses of this paper (*caveat emptor*).

Brune relies heavily on a particularly simple solution for the displacements that occur due to a step in shear traction on a free surface. This problem was described back in Chapter 3 (see 3.58). In particular, if an instantaneous change in shear traction, $T_x = \sigma_e H(t)$ occurs uniformly on the surface of a uniform half space, where the plane is described by $z = 0$, then a uniform shear stress propagates into the medium (in the z direction) at the shear-wave velocity. That is, the solution is just $\sigma_{xz} = \sigma_e H\left(t - \dfrac{z}{\beta}\right)$. This means that

$\varepsilon_{xz} = \dfrac{\sigma_e}{\mu} H\left(t - \dfrac{z}{\beta}\right) = \dfrac{\partial u_x}{\partial z}$. In equation (3.68), we derived that $\varepsilon_{xz} = \dfrac{1}{2}\dfrac{\dot{u}_x}{\beta}$. So

$\dot{u}_x = 2\beta\varepsilon_{xz} = \dfrac{2\beta}{\mu}\sigma_e H\left(t - \dfrac{z}{\beta}\right)$. This can be integrated to give the displacement,

$$u_x = \frac{2\beta}{\mu}\sigma_e\left(t - \frac{z}{\beta}\right)H\left(t - \frac{z}{\beta}\right) \qquad\qquad 8.58$$

This is a simple linear ramp in time that starts at $t = \dfrac{z}{\beta}$. The Fourier amplitude spectrum of this displacement can be easily deduced by recognizing that a linear ramp is the second time integral of an impulse. Now recall that integration in the time domain is the same as dividing by $i\omega = 2\pi i f$ in the frequency domain. Now since $\left|\tilde{\delta}(f)\right| = 1$,

$\left|\tilde{u}_x\right| = \dfrac{-2\beta}{\mu}\sigma_e\left(\dfrac{1}{2\pi f}\right)^2$. **That is, a simultaneous step in stress on the fault should produce ground displacement with an f^{-2} spectrum.** Brune argued that the accelerations experienced adjacent to the rupture were a measure of the effective stress accelerating the sides of the fault.

The solution that Brune used **assumes uniform stress applied instantaneously on an infinitely large fault**, and the resulting fault slip increases linearly with time indefinitely. That is, the slip trends to infinity. Brune reasoned that the final slip should approach the slip expected from fault dimensions and the appropriate stress drop (similar to Aki's argument). He hypothesized that, in the very near-source region,

$$u_x \approx \frac{\sigma_e}{\mu}\beta\tau\left(1 - e^{\frac{t}{\tau}}\right)H\left(t - \frac{z}{\beta}\right) \qquad\qquad 8.59$$

where $\tau = \dfrac{r}{\beta}$, and r is the radius of an approximately circular rupture area. This leads to particle velocities given by

$$\dot{u}_x = \frac{\sigma_e}{\mu} \beta e^{-1/\tau} H\left(t - \frac{z}{\beta}\right) \qquad\qquad 8.60$$

The Fourier amplitude spectrum of 8.59 is

$$|\tilde{u}_x| = \frac{\sigma_e}{\mu} \beta \frac{1}{\left(1 + i2\pi \dfrac{f}{f_c}\right)^2} \qquad\qquad 8.61$$

As before, this spectrum is f^{-2} , **which is the inevitable result of the discontinuous jump in particle velocity at the start of the motion**. Brune argues that this jump is fundamental to the dynamics of earthquakes. Furthermore, he argued that the dynamic stress controls the particle accelerations in the very near-source region. Notice that the particle acceleration that is derived by differentiating 8.60 is

$$\ddot{u}_x = \frac{\sigma_e}{\mu} \beta e^{-t - z/\beta/\tau} \left[\left(\frac{1}{\tau}\right) H\left(t - \frac{z}{\beta}\right) + \delta\left(t - \frac{z}{\beta}\right)\right] \qquad\qquad 8.62$$

The **impulse function comes from differentiation of the step function. This indicates infinite accelerations at the arrival time of the shear wave. Brune argued that structural heterogeneity scattered these high-frequency waves and that the result is random white noise whose amplitude is determined by the effective stress**.

For frequencies between $f = 0$ and $f = f_c$, Brune's spectrum 8.61 is similar to the Aki spectral model. Brune's argument about instantaneous stress drop led him to conclude that the high-frequency decay was naturally f^{-2} and that the accelerations had a white-noise spectrum whose amplitude scaled linearly with the effective stress. When Brune imagined this model, he was aware that earthquake engineers were simulating recorded near-source accelerograms as Gaussian white noise modulated by an envelope function in time. **Brune then appealed to conservation of radiated energy (through different enclosing spheres) to conclude that if the high-frequency S-wave is f^{-2} in the near-source region, then it must also be f^{-2} in distances that are large compared to the source dimension**.

At this point, Brune had used inductive reasoning to describe a Fourier amplitude spectrum for a distant observer. This spectrum had a long-period amplitude that is scaled by the potency (he used moment), a corner frequency that described the dimension of the rupture (used to determine stress drop), and an f^{-2} high-frequency radiation whose overall amplitude is controlled by the effective stress. In his 1968 paper, Brune stitched these different features together to hypothesize that the displacements in seismograms observed at distances large compared to the rupture dimension should be of the form

$$u \sim f\left(\frac{r}{R}\right) \frac{\sigma}{\mu} \beta t \exp(t\alpha) H(t) \qquad\qquad 8.63$$

and that the distant spectral scaling should be

$$|\tilde{u}(f)| \approx \mathcal{R}_{\theta\varphi} \frac{\sigma_e \beta}{\mu} \frac{r}{R} F(\varepsilon) \frac{1}{4\pi^2 f^2 + \alpha^2} \qquad\qquad 8.64$$

where $\mathcal{R}$ is a radiation pattern, r is the radius of the rupture area, R is the distance of the observer, $\alpha = 2.21\beta/r$ (the corner frequency). and $\varepsilon \equiv \dfrac{\Delta\sigma}{\sigma_e}$, which Brune called the fraction of the stress drop. $F(\varepsilon) = \left\{ (2-2\varepsilon)\left[1 - \cos\left(2.42\pi\varepsilon f/\alpha\right)\right] + \varepsilon^2 \right\}^{\frac{1}{2}}$; this factor joins the high-frequency levels to the corner with a section of the spectrum that decays as f^{-1} . The effect of partial stress drop ($\varepsilon < 1$) is shown in figure 8.8.
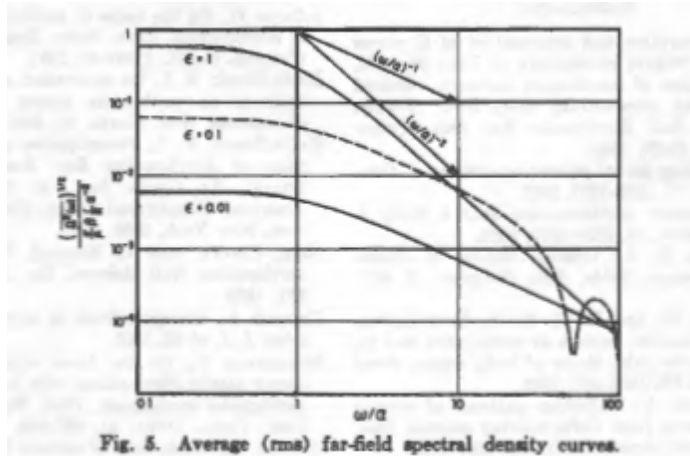

Fig. 5. Average (rms) far-field spectral density curves.

Figure 8.8. From Brune (1970). This is the "Brune far-field spectral model." If the stress drop is the effective stress then it's an f^{-2} with a single corner. If the stress drop is less than the effective stress, then there are both f^{-2} and f^{-1} falloffs with two corners.

Although Brune introduced the concept of partial stress drop, it is very rarely used, and most researchers use a simpler version of the Brune spectrum as given by

$$\left| \tilde{u}(f) \right| \sim \frac{P}{1 + \left(\dfrac{f}{f_c}\right)^2} \tag{8.65}$$

This simplified form assumes that the effective stress (which regulates high-frequency radiation) equals the final change in stress (which is controlled by the rupture dimension and the Potency). This simplified form of the model links the high-frequency near-source radiated energy to the overall size of the earthquake (P) and the stress drop. This can be seen in the following argument, which I developed in 2013.

Alternate Derivation of Brune Spectral Scaling.

Although I read Brune's paper many times, I find that its logic is difficult to follow. To better understand the "essence" of the Brune model, I devised an alternate derivation of his spectral law for waves observed at distances larger than the rupture distance. My goal is to start the derivation with a minimum number of simple assumptions. The

following is my derivation. It begins with several key assumptions that may be stated somewhat differently than Brune did.

1) **Assume that the duration of the S-wave group is proportional to the dimension of the fault** (that is, assume constant rupture velocity), this assumption can be restated $f_c \sim \dfrac{1}{V_R\sqrt{S}}$. Brune also made this assumption.

2) **Assume size similarity of the form** $P \sim S\bar{D} \sim S^{3/2}\Delta\sigma$, **or alternatively** $LW \sim P^{2/3}\Delta\sigma^{-2/3}$. Brune also made this assumption.

3) **Assume that the Fourier amplitude spectrum is a constant at low frequencies and f^{-2} at high frequencies as shown by equation** 8.65.

Assumptions 1 and 2 imply that $f_c \sim P^{-1/3}\Delta\sigma^{1/3}$. Substituting into 8.65, we obtain

$$\left|\tilde{u}(f)\right| \sim \frac{P}{1+\left(\dfrac{f}{P^{-1/3}\Delta\sigma^{1/3}}\right)^{2}} \qquad\qquad 8.66$$

Which has asymptotes

$$\left|\tilde{u}(f)\right| \sim \begin{cases} P & f \ll f_c \\ P^{1/3}\Delta\sigma^{2/3} f^{-2} & f \gg f_c \end{cases} \qquad\qquad 8.67$$

This is an alternate form of Brune's spectrum that is written as an explicit function of stress drop. Variations of Equation 8.66 are so often used (aka, assumed) that I will refer to this as **the Standard Source Model.** Now the radiated energy spectrum scales as the square of the Fourier amplitude spectrum, or at high frequencies

$$E_R^{f\gg f_c} \sim \left(\tilde{U}_R^{f\gg f_c}\right)^2 \sim P^{2/3}\Delta\sigma^{4/3} \sim \left(S^{3/2}\Delta\sigma\right)^{2/3}\Delta\sigma^{4/3} = S\Delta\sigma^2 \quad 8.68$$

That is, **Brune's spectral model predicts that high-frequency radiated energy scales with the rupture area and the square of the stress drop.** This is consistent with the notion that the amplitude of very high frequencies close to the rupture are a constant that is related to the effective stress. This is a key feature of Brune's spectrum; the near-source acceleration should scale with stress drop. That is, **at very close distances, pga should be independent of the earthquake magnitude**, but it should scale linearly with stress drop. As I will shortly show, near-source pga is indeed independent of magnitude (M > 6). However, **near-source accelerograms show that near-source pga is also independent of stress drop. This turns out to be a fundamental flaw of the Standard Source model**.

Stress Drop and Near-Source Ground Motions

I will now discuss the scaling of near-source ground motions and how they scale with magnitude and stress drop. The work in this section is modified from an unpublished manuscript by Tom Heaton and Masumi Yamada. This is work from 2013.

Brune (1970) hypothesized that near-source pga correlates linearly with effective stress, σ_e. Since it's not currently feasible to measure σ_e independently from measures of radiated seismic waves, it is not feasible to independently verify Brune's hypothesis. It is, however, possible to investigate the correlation between near-source motions (less than 10 km from the surface projection of the fault rupture) and spatially averaged static stress drop, $\overline{\Delta\sigma}$; this is a meaningful test since the most common form of Brune's model assumes that $\Delta\sigma = \sigma_e$. In this study, we show that there is no significant correlation between near-source peak ground accelerations (pga's) and spatially averaged static stress drop $\overline{\Delta\sigma}$. In contrast, near-source long-period motions do correlate with stress drop.

Our current study investigates the relationship between static stress drop and the intensity of near-source ground motion as measured by peak ground acceleration (pga), peak ground velocity (pgv), and peak ground displacement (pgd). This work is a follow-on to our previous study of the statistical characteristics of near-source motions, where we reported that there is no apparent correlation between near-source pga and pgd (Yamada, Olsen, and Heaton, 2009, *Statistical Features of Short-Period and Long-Period Near-Source Ground Motions,* Bulletin of the Seismological Society of America, 99 (6). pp. 3264-3274. ISSN 0037-1106). In this earlier study, we hypothesized that pgd should correlate with stress drop; if all other source parameters are equal, then larger slips are associated with both larger near-source pgd's and also with larger stress drops. However, the lack of correlation between near-source pga and pgd seems to imply that pga and stress drop are also uncorrelated.

In this study, we estimate the average static stress drop for 20 earthquakes for which there are also near-source strong motion accelerograms. We then investigate the relationship between measures of shaking intensity and earthquake potency and stress drop.

We estimate average static stress drop using the relations between average slip, $\overline{D}$, rupture length, L, and rupture width, W, for a rectangular fault in a homogeneous half space as reported by Parsons et. al. (1988) and presented in equations 8.26 through 8.29. In particular, we assume that

$$\overline{\Delta\sigma} \approx \mu\left(C\frac{\overline{D}}{W} \right)$$
$$= \mu\left(C\frac{P}{LW^2} \right)$$

8.69

Where $P = LW\overline{D}$ is the Potency (the volume integral of inelastic shear strain in an earthquake), C is a dimensionless constant that depends on the aspect ratio of the rupture

$\dfrac{L}{W}$, the depth of burial of the rupture z_{top} , and rupture dip and rake angle.

Approximations for *C* are given in the section on stress drop.

For ruptures with shallow burial depth, C changes quickly as a function of $\left(z_{top}\middle/W\right)$; C

is within 95% of its value for deeply buried when $\left(z_{top}\middle/W\right) = 0.16$. This rapid change in

C with burial depth is caused by the fact that, in an elastic model, there are very large strains and stresses in the region between the top of the fault and the free surface. In reality, it is highly questionable whether such large stresses develop in the shallow region just above a rupture that does not quite reach the surface. For simplicity, we classified events as either "surface rupturing" or "deeply buried"; that is, we assumed that C is described by 8.29 for any earthquake that did not have surface rupture. We assume a uniform rigidity of $\mu = 35\,\text{GPa}$, which is taken to represent average properties in the upper Crust.

Table 8.1 lists the source parameters for the earthquakes in this study. Most of these models are described in the ETH source data base. For some of the events there are multiple source models and, in these cases, we assume that the models are log-normally distributed about the geometric mean of the models. Figure 8.10 shows the relationship between average static stress drop and Moment Magnitude, $\mathbf{M} \equiv 1 + \frac{2}{3}\,LogP$. In order to show the distribution of the different models, each of the earthquakes is assigned a single $\mathbf{M}$ that is based on the geometric mean of the Potencies of the different models. In a similar manner, the static stress drop values for each event are the geometric mean of the static stress drops for multiple models for each earthquake. Of course, it's difficult to estimate the uncertainty in these values (especially stress drop), but we have included an error estimate that is based on the standard deviation of all of the models about their mean. We calculate the standard deviation using Bessel's correction for small number statistics. Earthquakes with only one source model are considered to be more uncertain than those that represent the geometric mean of multiple models

Earthquake	Year	Date	NS	mean(Mw)	mean(Δσ)	std(Δσ)
Imperial Valley	1979	10/15	12	6.53	1.00	0.33
Loma Prieta	1989	10/17	6	6.99	5.47	1.27
Landers	1992	6/28	1	7.25	2.89	1.00
Northridge	1994	1/17	7	6.79	3.17	1.15
Kobe	1995	1/17	4	6.96	1.25	0.50
Izmit	1999	8/17	3	7.51	2.01	0.88
Chi-Chi	1999	9/20	38	7.68	2.39	0.47
Western Tottori	2000	10/6	5	6.84	2.11	0.42
Denali	2002	11/3	1	7.91	2.97	1.66
Parkfield	2004	9/28	48	6.05	0.15	0.02
Mid Niigata	2004	10/23	6	6.72	0.97	0.00
Noto-Hanto	2007	3/25	1	6.74	1.63	0.00
Niigataken-Chuetsuoki	2007	7/16	2	6.81	2.17	0.00
Wenchuan	2008	5/12	6	8.00	1.67	0.00
Iwate-Miyagi	2008	6/14	6	6.96	2.43	0.00
Surugawan	2009	8/11	1	6.43	1.19	0.00
Darfield	2010	9/4	7	7.00	0.83	0.00
Christchurch	2011	2/22	15	6.36	2.83	0.00
Northern Nagano	2011	3/12	4	6.35	1.92	1.08
Fukushima-Hamadori	2011	4/11	3	6.69	1.13	0.00
Total			176			

Table 8.1. Earthquakes used in this study have finite source models tabulated by the SRCMOD project (Mai, ETH). NS is the number of stations. All statistics are computed assuming log-normal distributions and Bessel's correction for small samples is assumed to calculate the standard deviation of $\overline{\Delta\sigma}$.
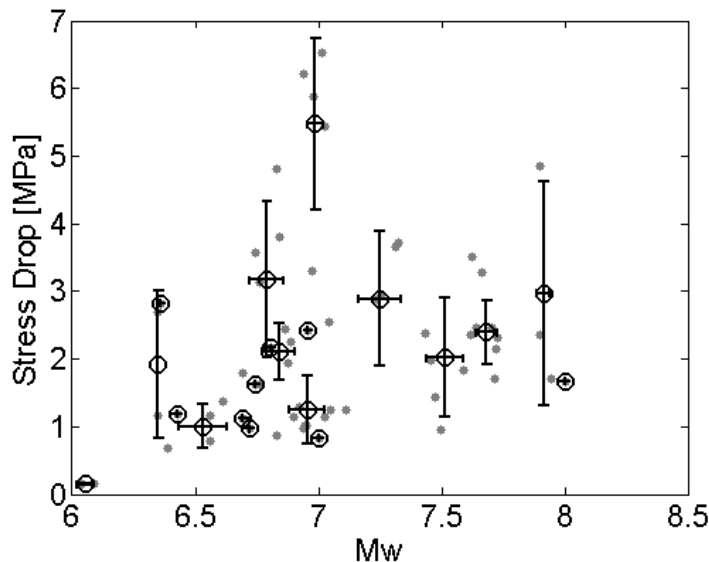

Figure 8.9. The values of $\overline{\Delta\sigma}$ plotted as a function of $\mathbf{M}$. Each dot represents a finite-source model compiled in the ETH database of finite source models. The circled points represent the mean model for a given earthquake. The "error bars" are the standard deviations for earthquakes that have more than one model.

We collected strong motion records from twenty earthquakes in table 8.1 from sites in the near-source region. We define the near-source region to be within 10 km of the surface projection of the fault rupture (often called the Joyner-Boore distance). We choose this number because it is small enough to demonstrate the source scaling, but large enough to include a sufficient number of records to characterize the statistics. To determine the PGA, PGV and PGD of a recorded ground motion, we processed each time history. We first remove the bias from the acceleration record by subtracting the mean. Then we integrate once and remove the long-period component with a fourth-order high-pass Butterworth filter (13.3 sec corner period) to generate the velocity time history. The displacement records are obtained from the integration of velocity waveforms. Peak values are obtained from the square root of the sum of the squares of the north-south, east-west and up-down components at each time step of the record.

Figure 8.10 shows a log-log plot of near-source pga as a function of static stress drop. Figure 8.10 also shows a linear regression of log pga as a function of $\mathbf{M}$ and log $\Delta\sigma$. Compatible with the conjecture of Yamada and others (2009), we do not observe a significant correlation between near-source pga and either stress drop or Potency (Moment Magnitude). Notice that there is no magnitude dependence to the near-source pga. This is because we constrained the magnitude dependence to be positive only. Allowing decreasing near-source pga with increasing $\mathbf{M}$ actually provides a marginally better fit to the data, but including this possibility requires us to propose non-intuitive models that only fit highly scattered data marginally better.
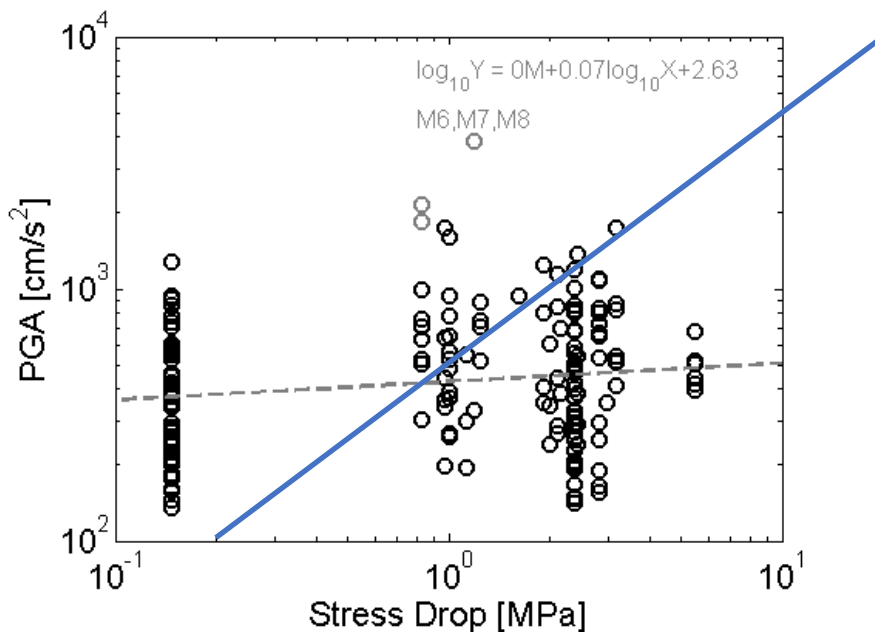

Figure 8.10. Near-source pga vs stress drop for all records used in this study. The dotted line represents the least-squares linear regression, $\log(pga) = 0.0M + 0.07\log\Delta\sigma + 2.63$. This seems to indicate that near-source pga is independent of magnitude (as Brune hypothesized) and also independent of stress drop (very different from Brune's hypothesis which is represented by the blue diagonal line).

Figures 8.11 and 8.12 are similar, but they are for near-source pgv and pgd, respectively.
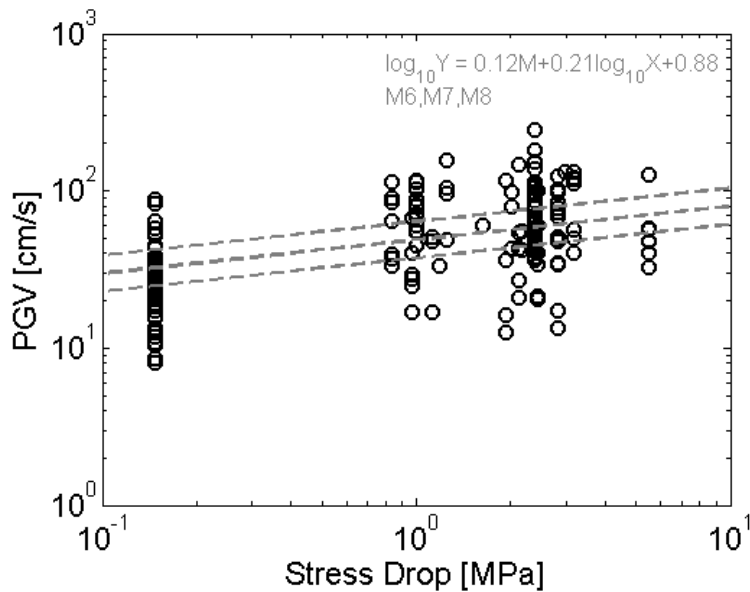

Figure 8.11. Near-source pgv vs stress drop for all records used in this study. The three dotted lines represent M 6, 7, and 8 in the linear least-squares regression $\log\left(pgv\right) = 0.12M + 0.21\log\Delta\sigma + 0.88$
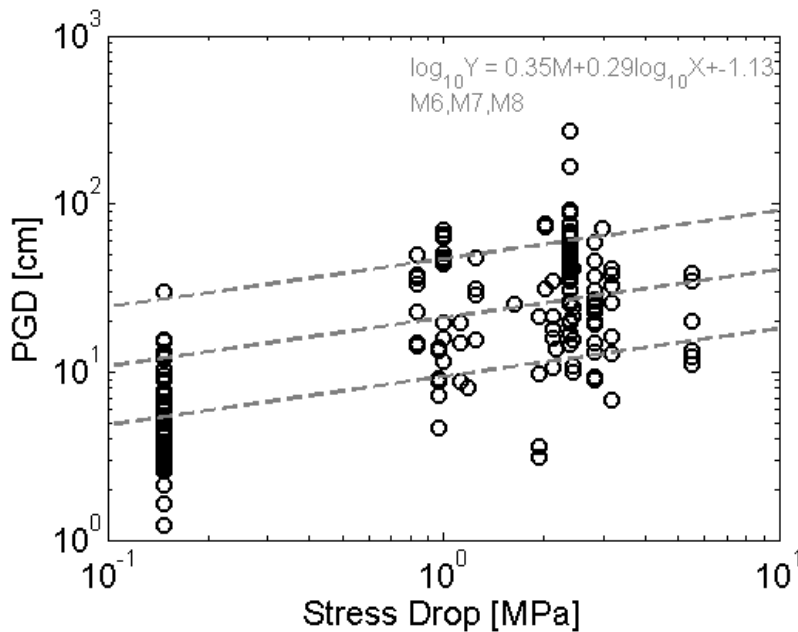

Figure 8.12. Near-source pgd vs stress drop for all records used in this study. The pgd is from high-pass filtered records with a 13 s corner. The three dotted lines represent M 6, 7, and 8 in the linear least-squares regression $\log\left(pgd\right) = 0.35M + 0.29\log\Delta\sigma - 1.13$.

We can anticipate the following asymptotic behavior for any ground motion prediction equations. When the distance is large compared to the source dimension, and when the predominant periods of the ground motion are large

compared to the source duration, we expect the ground motion amplitudes to scale with the seismic potency, or

$$\log U_{far\&lowfreq} \sim \log P \sim \frac{3}{2} M \qquad 8.70$$

For very near-source long-periods (e.g. displacement), we expect the peak amplitude to scale with the size of the slip on the nearby fault segment, or

$$\bar{D} = \frac{P}{S} \qquad 8.71$$

and

$$S = \left(\frac{P}{\Delta\sigma} \right)^{\frac{2}{3}} \qquad 8.72$$

So

$$D \sim P^{\frac{1}{3}} \Delta\sigma^{\frac{2}{3}} \qquad 8.73$$

So

$$\log U_{near\&lowfreq} \sim \log \bar{D} \sim \log P^{\frac{1}{3}} + \log \Delta\sigma^{\frac{2}{3}}$$
$$\sim \frac{1}{2} M + \frac{2}{3} \log \Delta\sigma \qquad 8.74$$

Relationship 8.74 can be compared with Figure 5 that is derived from the ground motions and stress drop models

$$\log pgd_{near} \sim 0.35M + 0.29 \log \Delta\sigma \qquad 8.75$$

This can be compared with the 10-s response spectral acceleration, sa_{10}, scaling reported by Campbell and Borzorgnia (2014).

$$Log\left(sa_{10} \right) \sim 0.333M \qquad 8.76$$

Cua and Heaton (2009) also produced a gmpe for pgd using data from both small- and large-magnitude earthquakes (see Figure 8.18). Their near-source pgd scales as

$$\log\left(pgd \right) \sim \frac{1}{3} M \qquad 8.77$$

Interestingly, Campbell and Borzorgnia (2007, PEER report) recognized that sa_{10} should approximately mimic the scaling of pgd, and that near-source pgd is controlled by average slip, which in turn, scales with $\Delta\sigma$. However, they state that "the resulting PGD ground motion model is intended for evaluation purposes only at this time and should not be used for engineering design until further empirical and theoretical verification becomes available." That is, the requirement that long-period near-source motions should scale with slip somehow seems to have never made it into the NGA2 gmpe's that were used to calculate the National Probabilistic Hazard Models. **It seems quite clear that $\Delta\sigma$ has a 1st order effect on near-source long-period motions. $\Delta\sigma$ should be**

included in the gmpe's for long-period motions. The current practice of only using magnitude and observer distance to predict long-period motions is flawed and is likely to seriously under predict the maximum motions from a suite of plausible events.

As an example of how these relations work, consider the implications for pga and pgd for a large strike-slip earthquake similar to the 1906 San Francisco earthquake. This event has been assigned a moment magnitude of M=7.8. If we were to assume that the rupture dimensions in a future earthquake were the same, but that the slip in the event was doubled, then the Potency and the stress drop would both double. Doubling the potency increases the moment magnitude by 0.2 units and inserting these values into 8.75 gives a 0.3 unit increase in log u, which corresponds to a doubling of u.

Figure 8.13 clearly shows how differently near-source pga scales with size than does pgd. This figure is from Yamada, Heaton, and Olsen (2009) and it shows a log-log plot of pga vs. pgd for every strong motion record for earthquakes of M>6 that was available at the time of the study. The light grey points are for sites at JB distances > 10 km, whereas the bold circles are for JB distances<10 km. Although there is a lot of scatter in in the relationship between pga and pgd, it is obvious that pga and pgd have a linear correlation for distances>10 km, whereas there is absolutely no correlation for distances < 10 km. This observation is compatible with the hypothesis that near-source pga is unrelated to stress drop.

Figure 8.14 is also from Yamada and other (2009) and it shows the frequency distribution of near-source pga. This figure clearly shows that near-source pga for M > 6 is approximately log normal with a geometric mean of 4.64 m/s2. The grey line shows the distribution without including a large number of records recorded in the 1999 M 7.6 Chi-Chi, Taiwan earthquake, whereas the bold solid line includes Chi-Chi data. Notice that the distribution of near-source pga for smaller previous events seems to fit the data that was observed in Chi-Chi. Again, this is convincing evidence that near-source pga completely saturates with magnitudes greater than 6.
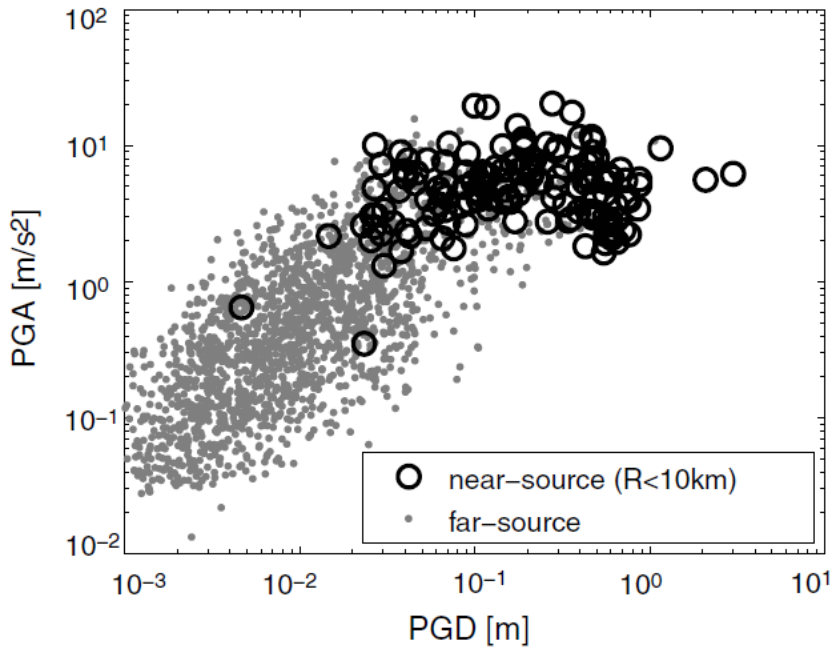
Figure 8.13. Pga vs. pgd for all strong motion records available as of 2008 (from Yamada, Heaton, and Olsen, 2009.
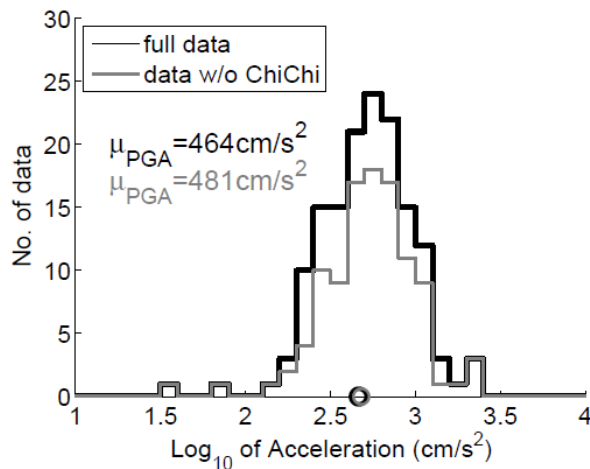

Fig. 8.14. Histogram of the number of near-source records as a function of log pga for earthquakes M.6. The figure is also from Yamada, Heaton, and Olsen (2009).

The saturation of near-source pga can also be clearly seen in the ground motion prediction equations (gmpe's) of Cua and Heaton (2009, *Characterizing Average Properties of Southern California Ground Motion Amplitudes and Envelopes.* Earthquake Engineering Research Laboratory, Earthquake Engineering Research Laboratory , Pasadena, CA. EERL report 2009-05). These gmpe's were created for use in earthquake early warning systems that require the ability to predict over a very wide range of magnitudes. Most gmpe's in engineering seismology are intended for use in predicting damaging motions (M > 5). The Cua and Heaton gmpe was created using a large data set recorded by the Southern California Seismic Network. In addition, strong motion data compiled by PEER

was also included. The solid lines are gmpe's derived using only sites on rock, whereas the dotted red line was derived using only data from soil sites.
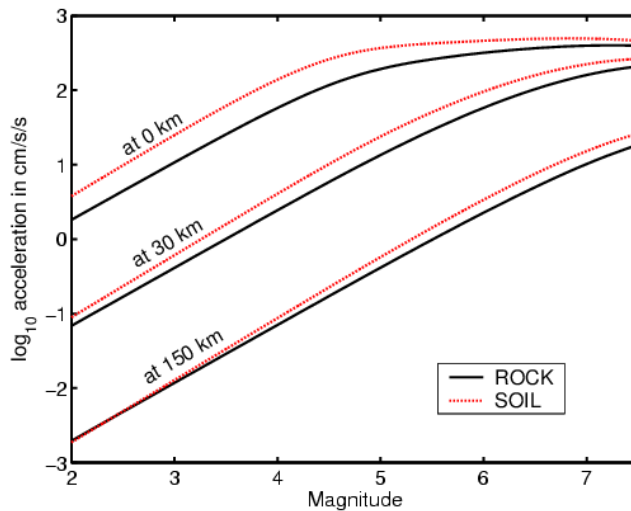

Figure 8.15. Relationship between pga and magnitude as observed at different Joyner-Boore site distances. From Cua and Heaton (2009)

Notice that the 0-distance gmpe seems to indicate that pga saturates for M>5 and the saturation level is consistent with Fig. 8.14. This saturation becomes less apparent for larger distances.   This is most likely caused by the fact that near-source pga is primarily the result of frequencies greater than 3 Hz, whereas the important frequencies become lower the further away that you are. Similar plots are shown for pgv and pgd in Figures 8.16 and 8.17, respectively. Notice that log(pgd) scales approximately linearly with magnitude for near-source and M < 5, whereas it scales as approximately $\frac{1}{2}M$ for M > 5; this is consistent with 8.75 . The linear scaling of log(pgd) with M seen at either smaller magnitudes or larger distances is consistent with 8.12.
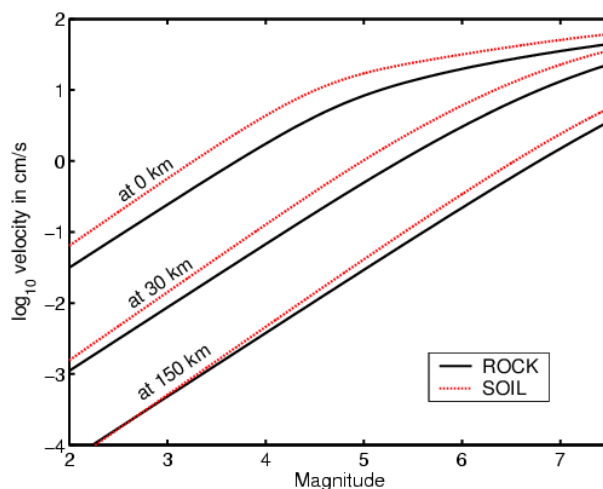

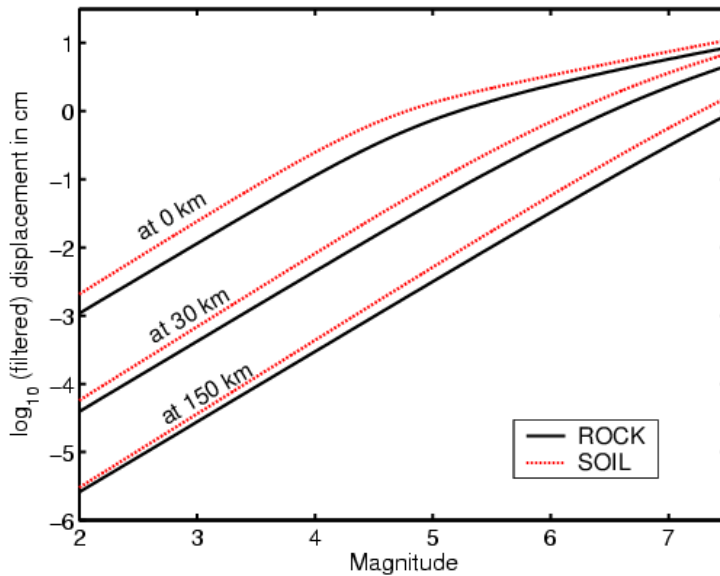Figure 8.16 Same as 8.15, except for pgv.
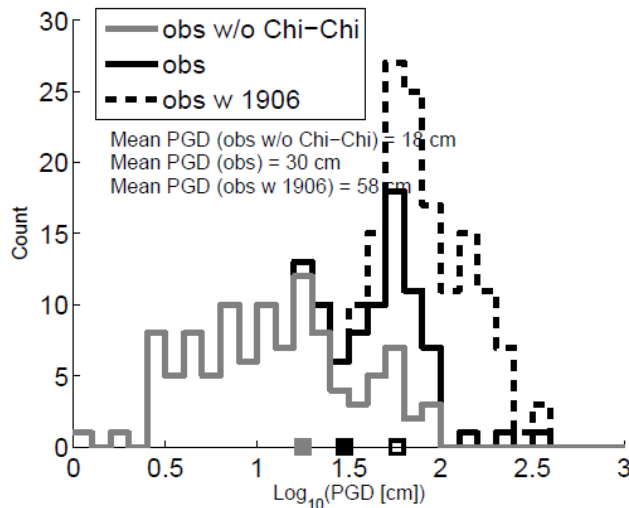
Figure 8.17 Same as 8.15, except for pgd.

Fig. 8.18. Same as Fig. 8.14, except for peak ground displacement (from Yamada, Olsen, and Heaton, 2009). The solid lines are the distributions with and without Chi-Chi data. The dotted line includes synthetic data at existing strong motion stations for a re-enactment of the 1906 San Francisco earthquake. The analysis in equation 8.24 suggest that this distribution may be log-uniform.

Figure 8.18 shows the number of strong motion records vs. log (pgd) as of 2009. Notice the contrast with log (pga) (see Fig. 8.15). Whereas the log (pga) data appeared to be log normally distributed about ½ g, the log(pgd) data is clearly not a normal distribution. As was argued in eq. 8.24, this distribution may be **log uniform**. The smaller pdg's are mostly from relatively frequent earthquakes of 6 < M < 7. Notice that including the data from just several large earthquakes (i.e., Chi-Chi and a re-enactment of 1906 San Francisco) can significantly alter this distribution, including the mean. Figure 8.18 is an example of a heavy-tailed power law distribution. Later I will present a more extensive development of the physics associated with probabilities that are defined by a power law. A common

feature of power-law distributed data is that the mean of the data typically increases as more data is added; this is certainly the case for the data set shown in Figure 8.18.

A log uniform distribution is an example of **Benford's law**, which is also sometimes referred to as the **first-digit law**. Benford's law was reported in 1938 and it is based on the observation that the first digit of numbers in tables of physical constants are often distributed logarithmically; numbers that start with the digit 1 are much more common than numbers that start with a 9. The formal definition of Benford's law is

$$P(N) = Log(N+1) - \log(N)$$
$$= \log\left(\frac{N+1}{N}\right) = \log\left(1 + \frac{1}{N}\right) \qquad 8.78$$

Where N is the first digit of the numbers in tables of physical measurements. To derive that this is, in fact, a power law, we can begin by assuming that $P(N) = \frac{1}{N}$, then the probability of digits between N and $N+1$ is

$$P(N) = \int_{N}^{N+1} \frac{dN}{N} = \int_{N}^{N+1} d(\log N) = \log\left(\frac{N+1}{N}\right) \qquad 8.79$$

which is Benford's law as given by 8.78.

It may seem surprising that pga, which is a measure of high-frequency motions, is more or less independent of both $\mathbf{M}$ and $\Delta\sigma$. However, as I will now show, saturation of high-frequency near-source motions with $\mathbf{M}$ is compatible with the well-known Brune (1970) spectral scaling model. *Unfortunately, it seems that Brune scaling is fundamentally inconsistent with the observation that near-source pga's that are independent of* $\Delta\sigma$. In the next section, I will show that it is not possible to formulate a simple spectral scaling law (single power law with one corner frequency) that has near-source pga that is independent of magnitude and which is also independent of stress drop.

In the variation of the Brune model where the static stress drop is assumed to be the effective stress, stress drop becomes the fundamental parameter that links statics and dynamics. However, many of the kinematic rupture models used in the Yamada and Heaton study actually reported slip pulse models. That is, the slip is already known at most points long before the rupture has stopped propagating. In slip pulse models, there is no correspondence between effective stress and stress drop. Two ruptures could be identical up to the point that one stopped and the other continued. The radiation from the ruptures would also be identical up to that point. However, since the rupture that continued would have a longer rupture length, it would have a different stress drop. **Stress drop is not a fundamental parameter in slip pulse models**; it's just a statistical accident that some ruptures continue on for long distances and have a lower stress drop.

In his 1970 paper, Brune assumed that the slip velocities (and accelerations) scale linearly with effective stress. He also hypothesized that near-source peak accelerations (pga) should scale linearly with effective stress. When using the Brune model, it is common to assume that effective stress is equal to stress drop. This logical construction is self-consistent only if 1) near-source pga's saturate with magnitude (they do), and 2) near-source pga's also scale linearly with stress drop (they seem to be independent of stress drop).

The conclusion of this discussion is that 1) the 1970 Brune model is incompatible with observed seismograms, and 2) there is no scale-free model similar to Aki's conjecture that fits available data.

Although Aki's conjecture is appealing, the most serious issue is that in order to maintain constant stress drop in crack-like ruptures, the fracture energy must grow as L (see 8.35). Clearly this violates the basic assumption that material properties are independent of scale.

Brune's spectral model has been an appealing and powerful conceptual framework to interpret seismic data. This framework consists of: 1) fit power-law envelopes to the Fourier amplitude spectrum , and 2) then derive the dynamic stresses in an earthquake. **Unfortunately, there are many major problems with Brune's approach.** I have already mentioned many of these problems in the previous discussion. I now gather these issues together in one list.

Problems with the Brune Spectral Model
I now list a number of specific objections that I have discovered. The first four objections are mistakes in continuum mechanics that are included in Brune's paper. The next five objections are descriptions of how earthquake observations are inconsistent with the Brune model.

1. *Near-field energy vs far-field energy*
   Perhaps the most fundamental error in the Brune paper is the use of the solution to an instantaneous traction applied to the surface of a half-space (equation 8.58) to infer f^{-2} spectral amplitudes for the far-field motion (equation 8.62). It's true that 8.58 is the solution to an instantaneous step in shear traction on the boundary of a half-space, but this solution is entirely comprised of near-field terms (see Chapter 7). In fact, the assumption of an infinitely large rupture surface means that there is no far-field radiation for this somewhat pathological problem. While it is true that energy is conserved for radiated far-field energy (except for anelastic attenuation), there is no such energy conservation between near-source ground motions (comprised of both near- and far-field terms) and far-field radiation (see Chapter 7). Also, recall from Chapter 7 that far-field terms are the time derivative of the near-field terms. That is, if the motion is a ramp in the near field (f^{-2}), then it is a step in the far field (f^{-1}).

2. *Infinite rupture velocity is impossible*
   Brune argued that "*the particle velocities are always much smaller than the rupture propagation velocity (since the stresses are much less than the shear modulus), and thus the interaction of particle motion velocity with the rupture velocity will always be small, i.e., the average energy spectrum for instantaneous application of stress will be approximately valid.*" Later analysis showed that the ratio of the rupture velocity to the shear wave speed is of interest to mode III ruptures, and the ratio of the rupture velocity to the Rayleigh wave speed is of interest for mode II ruptures. Since typical observations of rupture velocities are in the range of 0.8β to 0.9β, the approximation of $V_R \approx \infty$ is clearly inappropriate. In fact, rupture velocity plays a key role in almost all models of dynamic rupture.

3. *Ramp-like slip produces f^{-3} far-field radiation.*
   This objection is related to objection 1. In Chapter 7, I demonstrated that a line source with an instantaneous step in slip produces a rectangular far-field time function, which has a f^{-1} high-frequency spectrum. A rectangular source with an instantaneous slip has a far-field time function that is the convolution of two rectangles, which is a trapezoid that has a high-frequency decay of f^{-2} (one for timing associated with length and the other with width). A rectangular source with fault slip that is a linear ramp has a far-field time function that is the convolution of three rectangles (f^{-3}). The third rectangle is associated with Brune's slip function. The solution describing the displacements from an instantaneous change in stress on a circular fault (the problem Brune claimed to be solving) is actually a difficult mechanics problem and I am not aware of any closed-form solutions. Madariaga (1976, Dynamics of an Expanding Circular Fault, Bull Seisms Soc. Am, 66, 639-667) used finite elements to simulate this problem and he concluded that the far-field radiated S-waves have a high-frequency spectral decay of $f^{-2.5}$. In Madariaga's study, most of the high frequencies were radiated at the circumference of the rupture.

4. *The Brune model is a uniform stress drop crack model.*
   I have already mentioned this important issue several times. Constant stress drop cracks produce stress concentrations at the crack tips that grow linearly with the rupture dimension. These models are not scale independent.

5. *Slip is not ramp-like.*
   Ruptures that propagate at sub-shear rupture velocities produce slip functions that are similar to a time-transformed version of slip on a crack. The slip near the tip of a static shear crack increases as $\Delta\sigma K\sqrt{x}$, where x is the distance from the crack tip and K is a stress intensity factor. For a steadily propagating crack, x transforms to $x - V_R t$. This means that for any fixed point on the rupture, $D \sim \Delta\sigma K\sqrt{t}$. To obtain the far-field time function, you need to differentiate to

slip velocity, or $\dot{D} \sim \Delta\sigma K / \sqrt{t}$. Convolving with $1/\sqrt{t}$ is equivalent to multiplying with $f^{-\frac{1}{2}}$. This is sometimes called a **fractional integral**.

6. *Predicted slip durations much longer than observations.*
   The Brune model predicts that slip in the near-source region is given by 8.59, which predicts that slip continues until shear waves arrive from the outer edge of the rupture. That is, the slip at any point is comparable to the entire duration of an earthquake. As I will describe shortly, there is convincing evidence that the duration of slip at a point is short (less than 10%) compared to the overall rupture time; this is known as a slip-pulse model.

7. *Real earthquake far-field time functions are much more complex than Brune's model.*
   Brune's far-field time function (8.63) predicts a relatively smooth single pulse (see Figure 8.19). In contrast, observed far-field time functions are typically very complex. See the far-field time functions shown in Figure 8.7 as an example.
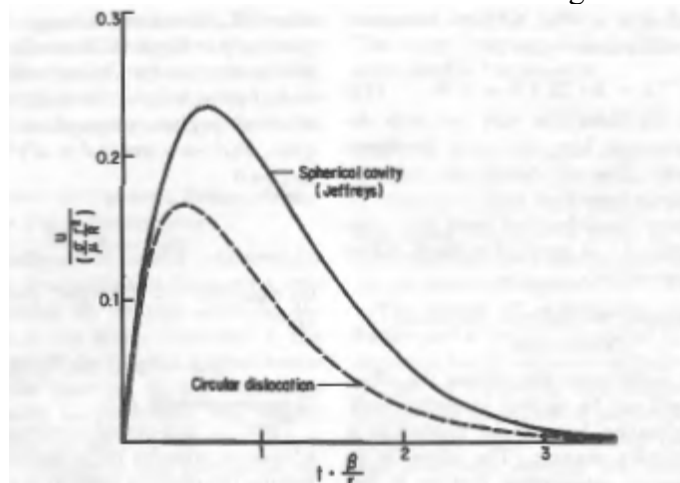

Figure 8.19 (from Brune, 1970) This is the time behavior
of Brune's far-field time function. It has the same ramp-
like beginning as his near-source ground motion and it
has a f^{-2} spectral decay.

Hartzell and Heaton (1093, Bull. Seism. Soc. Am. 73, 1553–1583) compiled far-field time functions from large earthquakes. More recently, Meier, Ampuero, and Heaton, (2017, *The hidden simplicity of subduction megathrust earthquakes.* Science, 357 (6357). pp. 1277-1281. ISSN 0036-8075) investigated characteristics of a compilation of time functions for large earthquakes and these are shown in Figure 8.20. The time functions are scaled to have the same total durations. The red curve at the bottom is the median of all of the time functions and it has a simple triangular form. However, the individual time functions are very complex and do not resemble Brune's time function.
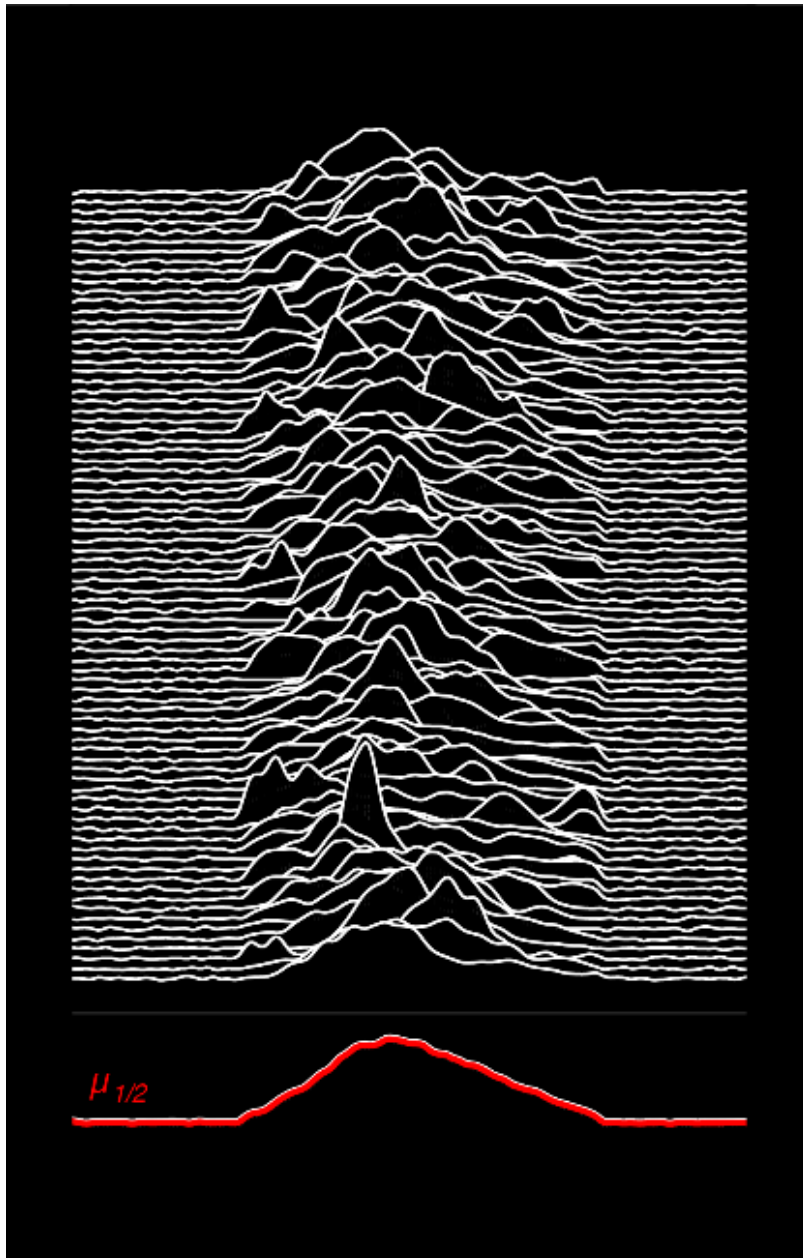
Figure 8.20. Teleseismic time functions from large subduction earthquakes. Notice the temporal complexity which contrasts with Brune's time function shown in Figure 8.7. The red line at the bottom is the median of all of the functions above it. This is an unpublished figure from the study of Meier, Ampuero, and Heaton (2017).

8. *Observed far-field spectral decays of large earthquakes (M > 6) are close to $f^{-1.5}$.* Although it is widely assumed that far-field radiated energy is characterized by f^{-2} spectral decay, recent studies using far more data than was available in the 1960's indicate that the average spectral decay between f_c and 0.5 Hz (the highest frequency observable at Teleseismic distances) is actually $f^{-1.5}$ as shown in Figure 8.21 from Hartzell and Heaton (1988, *Failure of self-similarity for large (Mw > 8 1/4)*

earthquakes. Bulletin of the Seismological Society of America, 78 (2). pp. 478-488. ISSN 0037-1106). Brune's conjecture that Fourier amplitude spectra of near-source high-frequency motions (frequency band from 2 Hz to 15 Hz) is approximately f^{-2} seems to be consistent with a very large volume of strong motion data. As I will discuss later, the physics of near-source radiation of high frequencies is still an unsolved problem and Brune's model is not a plausible explanation of the observation that near-source acceleration look like Gaussian white noise.
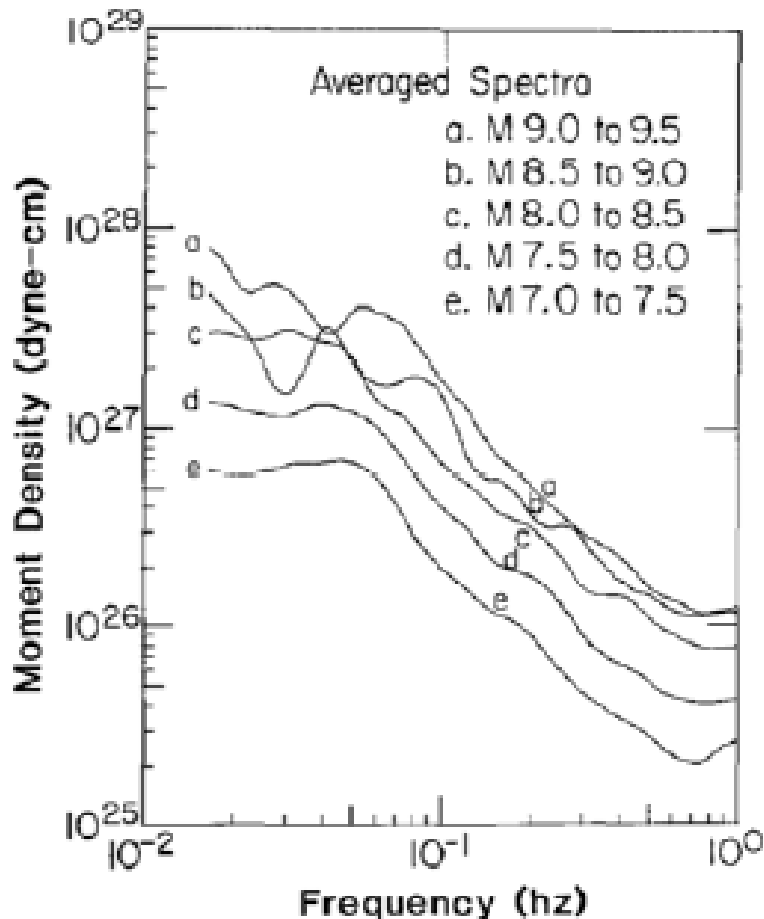

Figure 8.21. These are the Fourier amplitude spectra of Teleseismic P-waves from large subduction earthquakes. The average spectral decay is $f^{-1.5}$ for the frequency band between 0.5 Hz and 0.08 Hz. Anelastic attenuation prevents study at frequencies higher than 0.5 Hz using Teleseismic records. From Hartzell and Heaton (1988)

9.   *There is no correlation between stress drop and pga*
      Brune hypothesized that the spectral amplitude of near-source accelerations should scale linearly with effective stress, σ_e, which is equal to stress drop in the most commonly used form of Brune's model . This is an important issue since current predictions of strong shaking often use variations in regional stress drops to predict corresponding variations in high-frequency shaking. More fundamentally, I show in Appendix A that there is no possible power law with a

single corner frequency that produces bear-source pga's that both saturate with magnitude and which are also independent of stress drop.

Simulating ground motions using records from smaller earthquakes.

As much as I dislike the Standard Source model, it is simple and easy to use. For example, it allows us to use the record, $u_1(t)$, from a smaller potency magnitude M_1 earthquake to simulate the motion, $u_2(t)$, from a larger M_2 earthquake. The expected ratio of the amplitude spectra of the two events can be estimated from 8.66 .

$$
\frac{|\tilde{u}_2(f)|}{|\tilde{u}_1(f)|} \sim \frac{P_2}{\left(1+\left(\dfrac{f}{f_{c2}}\right)^2\right)} \left(\frac{1+\left(\dfrac{f}{f_{c1}}\right)^2}{P_1}\right)
$$

$$
= \left(\frac{10^{\frac{3}{2}M_2}}{1+\left(\dfrac{f}{f_{c2}}\right)^2}\right)\left(\frac{1+\left(\dfrac{f}{f_{c1}}\right)^2}{10^{\frac{3}{2}M_1}}\right)
$$

$$
= 10^{\frac{3}{2}(M_2-M_1)} \frac{1+\left(\dfrac{f}{f_{c1}}\right)^2}{1+\left(\dfrac{f}{f_{c2}}\right)^2}
$$

$$
= 10^{\frac{3}{2}(M_2-M_1)} \left(\frac{f_{c1}^2+f^2}{f_{c2}^2+f^2}\right)\left(\frac{f_{c2}^2}{f_{c1}^2}\right) \qquad\qquad 8.80
$$

Assuming stress drop invariance (that is $\Delta\sigma_1 \approx \Delta\sigma_2$), we can use 8.54 to estimate the corner frequencies of the two events. Or,

$$f_c \approx \left(\frac{V_R^3}{20\mu} \frac{\overline{\Delta\sigma}}{P} \right)^{1/3}$$

$$\approx \left(\frac{\left(2.8\times10^3\,m/s\right)^3 2.7\times10^6\,Pa}{7\times10^{10}\,Pa\times10^{3/2 M-1}\,m^3} \right)^{1/3}$$

$$= \left(\frac{2.2\times10^{10}\left(m/s\right)^3 2.7\times10^6}{7\times10^9\times10^{3/2 M}\,m^3} \right)^{1/3} \qquad 8.81$$

$$= \left(\frac{8.5\times10^7}{10^{3/2 M}\,s^3} \right)^{1/3} = 440\times10^{-1/2 M}$$

$$= 10^{2.64-1/2 M}\,s^{-1}$$

Substituting 8.81 into 8.80 gives

$$\frac{\left|\tilde{u}_2(f)\right|}{\left|\tilde{u}_1(f)\right|} \approx 10^{3/2(M_2-M_1)}\left(\frac{f_{c1}^2+f^2}{f_{c2}^2+f^2} \right)\left(\frac{f_{c2}^2}{f_{c1}^2} \right)$$

$$\approx 10^{3/2(M_2-M_1)}\left(\frac{10^{3.3-M_1}+f^2}{10^{3.3-M_2}+f^2} \right)\frac{10^{3.3-M_2}}{10^{3.3-M_1}} \qquad 8.82$$

$$= 10^{1/2(M_2-M_1)}\left(\frac{10^{3.3-M_1}+f^2}{10^{3.3-M_2}+f^2} \right)$$

Therefore, if we have recorded the motion $u_1(t)$ for a magnitude M_1, then we can obtain an estimate of the motion expected at the same site from a M_2 event at the same source location using the following approximation

$$u_2(t) \approx 10^{1/2(M_2-M_1)} FT^{-1}\left[\tilde{u}_1(f)\left(\frac{10^{3.3-M_1}+f^2}{10^{3.3-M_2}+f^2} \right) \right] \qquad 8.83$$

This approximation assumes that the observing distance is large compared to the source dimensions; that is, both events are approximately point sources. This requirement can be relaxed by assuming that a large rupture is approximately described by the appropriate sum of smaller ruptures. The responses of the smaller ruptures are assumed to describe the response of different parts of the finite fault. This technique is referred to as the **Empirical Green's Function** technique. I will not go through the details here (it's considerably more complex than what I've just derived), but it is discussed in detail in Appendix B of Heaton and Hartzell, 1989, Estimation of strong ground motions from hypothetical earthquakes on the Cascadia subduction zone, Pacific Northwest, Pure and Applied Geophysics, 129, 131-201.

Self-Healing Slip Pulses

Early in my research career, I investigated the physics of strong ground motions from significant earthquakes. In the 1970's this was a new field of study. Although there had already been strong motion recordings of several earthquakes, the 1971 M 6.7 San Fernando earthquake was the first earthquake with dozens of near-source strong motion records. Furthermore, there was an extensive project at Caltech to digitize the analog film accelerometer records and to integrate these records into ground velocity and displacement. One of the most significant records was from the Pacoima dam site that was directly up-dip from the hypocenter. This record had two key features of great interest to earthquake engineering (see Figure 8.22). The first was a peak acceleration of 1.15 g that occurred relatively late in the record. The second key feature was a distinct velocity pulse with an amplitude of 1.25 m/s. The acceleration records were too high-frequency to be interpreted using a deterministic model of rupture on a finite fault. However, the ground velocity pulse was very different. I interpreted it as a far-field S-wave that was strongly enhanced by directivity. Rather remarkably, the duration of this pulse was less than 1 s. Since the duration of the S-wave from slip on any point on the rupture is the duration of the slip at that point, the implication is that the duration of slip would need to be less than 1 second as well. When the rupture is propagating towards the receiver at close to the S-wave velocity, then all of the far-field S-waves arrive simultaneously, thereby producing this velocity pulse.
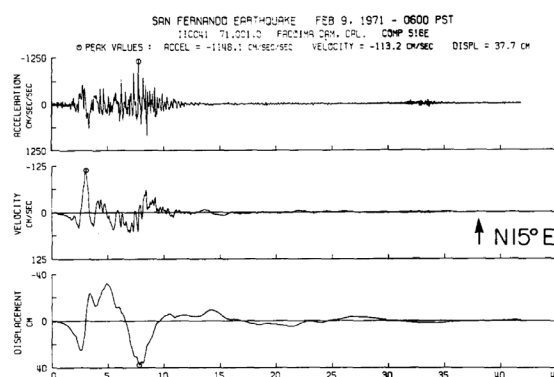

Figure 8.22. North component of ground motion from the Pacoima Dam recording of the 1971 San Fernando earthquake. The strong velocity pulse at 3 seconds is interpreted as a directivity enhanced far-field S-wave. Modeling of this pulse indicates that the duration of slip was less than 1 second at individual points on the 15-km length rupture, which took more than 5 seconds to occur.

One of my colleagues, Steven Hartzell, was also working in this same research direction and we collaborated on constructing finite-fault models of the ruptures to explain the near-source strong motion data from a number of earthquakes. All of these models shared the common feature that the duration of slip at any point was short compared to the time it took for the rupture to propagate over the rupture surface. For example, Figure 8.23 shows our preferred rupture model of

the M 6.5 1970 Imperial Valley Earthquake. The hatched region shows our
estimate of the part of the fault that was rupturing at a particular instant.
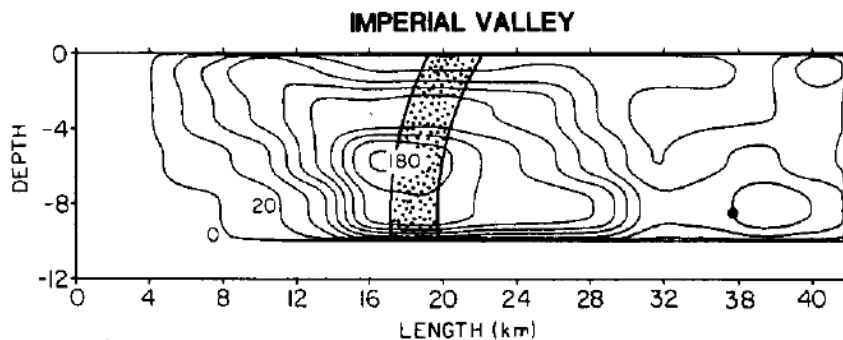

Fig. 8.23. Final slip distribution used to match strong
motion and teleseismic data for the 1979 Imperial Valley
Earthquake. The stippled region shows the area inferred to
be slipping at just one instance in time (from Heaton, 1990).

It was only later in 1988 that I realized that our rupture models were
fundamentally inconsistent with crack-like models, which were the prevailing
models for earthquake dynamics. In crack-like models, slip at a point starts when
the rupture arrives and then continues until S-waves arrive from the end of the
rupture.

Fig. 8.24 shows the slip histories for a uniform stress drop on a circular fault
embedded in a homogeneous whole space (from Madariaga, 1978). The rupture
expands radially from the center of the circle with a rupture velocity of 0.9β.
Notice that the slip at every point continues until a shear wave arrives from the
circumference of the fault (sometimes referred to as a stopping phase). This shear
wave contains the information that causes points on the fault to stop slipping. In
general, uniform stress drop ruptures in which the slip duration is determined by
the end of the rupture have slip durations that are on the order of $\dfrac{2\sqrt{S}}{3V_R}$.

These long slip durations contrast sharply with much shorter slip durations used
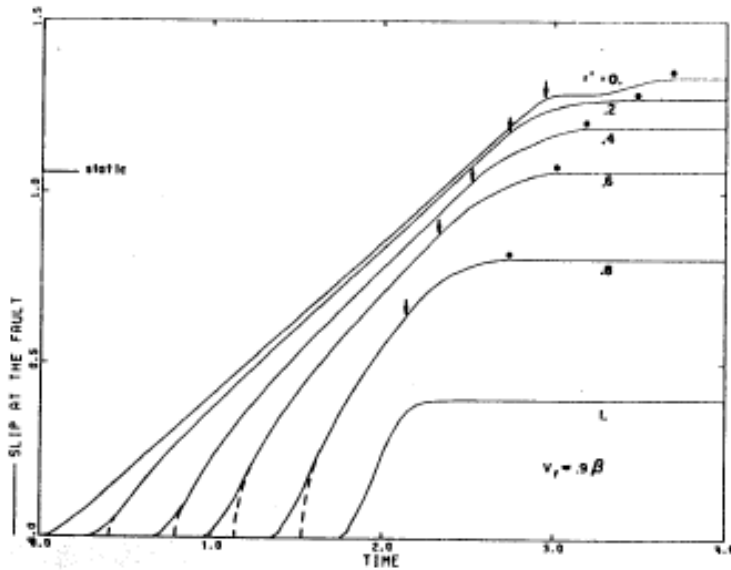to model seismic data from significant earthquakes.

Figure 8.23. Crack-like rupture. Slip histories for a variety of points on a circular fault in a homogeneous whole space. The upper diagonal is near the hypocenter (also the center of the circle) and the lower right history is near the outer edge of the circle. The stress drop is homogeneous and the rupture velocity is 0.9β . The arrows represent the time that a stopping phase comes from the perimeter of the rupture and the dots represent the cessation of rupture. (from Madariaga, 1976).

The primary question was to explain why slipping would stop before the rupture stopped propagating. This issue was the subject of my 1990 paper, Evidence for and implications of self-healing pulses of slip in earthquake rupture (Phys. Of Earth and Planet. Int., 64, 1-20). In addition to the source modeling of seismic records, there were also eyewitness reports of relatively short times for surface fault scarps to develop.

My paper was extremely controversial. In fact, I eventually withdrew it from the BSSA since the reviewers would not accept it without a wide range of fundamental changes. Since the time of the publication, there have been many other well recorded earthquakes and I think it's safe to say that it is now relatively well accepted that earthquake rupture durations at a point are small compared to the overall duration of the rupture.

I coined the phrase self-healing pulses of slip. I suggested that dynamic friction was strongly rate weakening. I argued that, similar to crack models, there is very high shear stress at the propagating crack tip and that this led to very high slip velocities just behind the propagating tip of the pulse. I suggested that the slip velocity decreased as $1/\sqrt{x_p}$, where x_p is the distance from the leading edge of the propagating pulse (see Figure 8.24). I hypothesized that the shear stresses in the vicinity of the crack tip were comparable to Byerlee friction and that the dynamic friction was very low in the high-slip velocity region just behind the rupture front. I hypothesized that as the slip velocity decreased with distance

8-49

from the crack tip, the dynamic friction increased. I suggested that this increase in dynamic friction caused the slip to arrest.

At the time of my paper, one of the most robust observations of earthquake scaling was the observation that average slip scales linearly with the rupture dimension (e.g., see Figure 8.5). In the slip-pulse model, slip at most points on the fault was already complete long before the final dimension of the rupture was determined. I suggested that there was a causal relationship between the amplitude of a slip pulse and the distance that the pulse was likely to propagate. That is, large earthquakes with long ruptures should have large slips that show up rapidly after the origin of the event. However, further study has shown that the characteristics of the early part of ruptures is not systematically different for larger ruptures.

I have now abandoned the hypothesis that larger earthquakes start with larger slips, and I now favor the notion that the correlation between slip and rupture length comes from the statistical properties of suites of spatially complex slip distributions. In order to adequately discuss slip pulses, it is necessary to discuss friction and the larger problem of maximum shear stress in the Earth's Crust. I will get to that shortly, but in the meantime, I have some comments about my hypothesis of strong rate weakening friction.

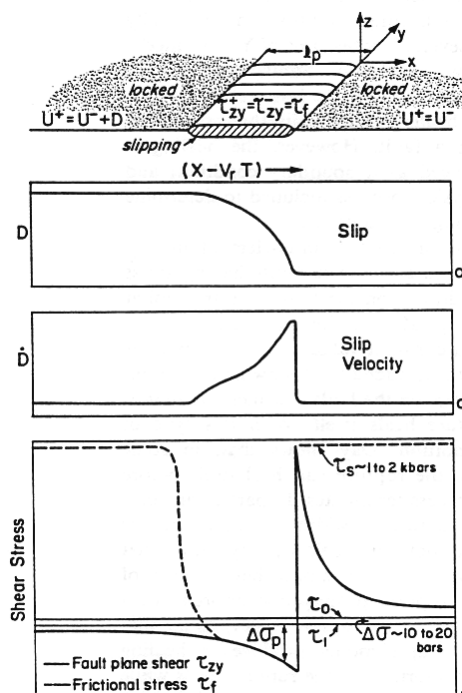Fig.8.24. Sketch from Heaton (1990) of a slip-pulse rupture mode. I proposed that fault friction decreased with slip rate. This implied that dynamic friction is low just after the passage of a rupture. The dotted line is meant to show the friction necessary for sliding. This friction is comparable to standard friction (60% of the confining stress) when the fault is not sliding. This model has instantaneous restrengthening.

When I first suggested strong rate weakening friction, I discussed the situation with SDSU Prof. Steven Day who was a leading expert in finite element simulations of dynamic shear rupture simulation. Steve was intrigued by the hypothesis and he attempted to simulate slip=pulse ruptures that arose from velocity weakening friction. He reported to me that he found it difficult to tune his friction and prestress in such a way that it would produce stably propagating slip pulses. In particular, they tended to either grow without bound as they propagated, or alternatively, they tended to die out fairly rapidly without propagating very far.

In retrospect, his finding was prescient, and should have been anticipated. That is, a velocity weakening friction law produces a positive-feedback dynamic system. In particular, the faster the sliding, the lower the friction. Lower friction produces faster sliding which leads to even lower friction. Conversely, slow sliding produces high friction, which causes even slower sliding. In fact, it has been shown that while strong velocity weakening can produce slip pulses, the pulses are inherently unstable. That is, there are no steady-state slip pulses caused by strong rate weakening friction (Elbanna, Lapusta, Heaton).

Since there are no stable steady-state solutions to strong rate-weakening slip pulses, they can only be studied using numerical techniques, such as finite elements or boundary-integral-elements. The fact that there are strong positive feedbacks in the dynamics means that the numerical solutions are sensitive to the details of the rupture dynamics, which means that high precision is required to make reliable numerical simulations.

Before I can discuss the dynamics of slip pulses, it is important to understand the characteristics of shear stress in the Earth's crust. I will shortly describe how understanding the characteristics of chaotic dynamics is key to understanding how it's possible to have 1) a static coefficient of friction of about 0.6, 2) low crustal heat flow and a paucity of frictional melts, 3) average static stress drops of about 1.5 MPa (1% of the overburden pressure), 4) radiated energy of about $2D \, \text{MJ} \big/ \text{m}_{\text{slip}}\text{-m}^2$.

ESTIMATING SHEAR STRESS IN THE CRUST

An enduring problem in geophysics focuses on estimating the size of shear stress that causes earthquakes. This problem has been approached in several ways. One approach is to estimate the stress necessary to cause shear failure. This approach typically leads to the conclusion that the frictional stresses at the depth of earthquakes are a substantial fraction of the pressure from the weight of the overburden rocks. As it turns out, these stresses far exceed any of the shear stresses that are estimated using seismic shaking data. This inconsistency has been debated for decades.

I have given numerous lectures about this issue and I often frame the problem as, "why are earthquakes so gentle?" Although this may at first sound like an odd

question, it is a rather serious issue when one is asked to estimate the maximum shaking that may occur in an earthquake.

In particular, laboratory test equipment that causes frictional sliding of surfaces that are forced together with the overburden pressure that occurs 10 km deep in the earth (about 200 MPa) are large and powerful machines. In the laboratory, yielding at these stresses can be dangerous even for relatively small samples; for safety, operators typically go to the next room to control these experiments. If you scale these experiments up to the size of significant earthquakes, then none of us would survive earthquakes, the ground motion would be too violent.

To better describe this problem, I will assume that earthquakes can be understood in terms of an expanding frictional crack. This will allow me to demonstrate fundamental inconsistencies that arise from this view. I will then present an alternate view that describes earthquakes as propagating slip pulses. Unfortunately, there are no analytic dynamic solutions to the slip pulse problem; by its nature, it is fundamentally a chaotic system.

A simple analysis of energy partitioning is appropriate at this point. Noting that an earthquake is the process of removing potential energy, E_W , from the crust and converting it into radiated seismic energy, E_R , frictional heat, E_f , and fracture energy, E_G. Conservation of energy dictates that

$$\Delta E_W = E_R + E_f + E_G \qquad 8.84$$

We can express the change in potential energy by integrating the traction on the fault over the slip, or

$$\Delta E_W = \iint_S \left(\int_0^{D(x,y)} \bar{\sigma}(x,y;D) dD \right) dxdy \qquad 8.85$$

Where $D(x,y)$ is the final slip as a function of position on the fault plane, and $\bar{\sigma} = \frac{1}{2}(\sigma_0 + \sigma_1)$ is the average shear stress between the initial state, $\sigma_0(x,y)$ and the final state $\sigma_1(x,y)$. Remember that the continuum is considered to be linearly elastic except on the slip surface. Therefore, the introduction of a reverse slip from the final state back to the original state would take the final stress back to the original stress. Since the system is linear, the average stress during that transition is the traction that the slip works against.

8.85 can be rewritten as

$$\Delta E_W = \iint_S \left(\int_0^{D(x,y)} \frac{1}{2} \left[\sigma_0(x,y) + \sigma_1(x,y) \right] dD \right) dxdy$$

$$= \iint_S \left(\int_0^{D(x,y)} \left[\sigma_0(x,y) - \frac{1}{2} \Delta \sigma(x,y) \right] dD \right) dxdy \qquad 8.86$$

Where $\Delta\sigma(x,y) \equiv \sigma_0(x,y) - \sigma_1(x,y)$. At this point, it is tempting to assume that the initial stress and the stress drop are approximately constant as a function of x and y. In fact, Kanamori and Heaton (2002) made this assumption. I will make it now because it helps to provide context for many other researcher's work. However, I warn you strongly that this is a *critical assumption that is almost certainly very wrong*. Because this is a key assumption (and probably wrong), I will introduce a notation for the energy estimated in this way. I will call it ΔE_W^{crack} , which I define as

$$\Delta E_W^{crack} \equiv \iint_S \left(\int_0^{D(x,y)} \left[\overline{\sigma_0} - \frac{1}{2}\overline{\Delta\sigma} \right] dD \right) dxdy \qquad 8.87$$

Where the overbars signify spatial average over the rupture surface. Stress drop and slip are both spatially smooth (except at the crack tip) for crack models, so the approximation is appropriate for cracks. In contrast, the slip in actual earthquakes is typically heterogeneous. Heterogeneous slip implies that prestress and stress drop are also spatially complex. Furthermore, the spatial complexity correlates between these quantities that specify the net energy change caused by an earthquake.

Examination of once-active faults that have been uplifted and exhumed by erosion shows that there is a paucity of frictional melting along rupture surfaces. This observation allows us to provide upper limits on the energy of frictional heating. In addition, seismic data allows us to estimate the radiated energy in earthquakes.

Note that $D(x,y)$ can be inferred from modeling seismic data. Knowing the spatial distribution of slip allow us to also infer $\overline{\Delta\sigma}(x,y)$. Now if we assume that slip amplitude is spatially smooth for crack-like ruptures, then $D^{crack}(x,y) \approx \overline{D}$; note that the averaging in this case is spatial averaging over the rupture surface . Using these assumptions allows us to integrate 8.87 to obtain

$$\Delta E_W^{crack} \approx \iint_S \left(\left[\overline{\sigma_0} - \frac{1}{2}\overline{\Delta\sigma} \right] \overline{D} \right) dxdy$$

$$\approx \left[\overline{\sigma_0} - \frac{1}{2}\overline{\Delta\sigma} \right] \overline{D}S \qquad 8.88$$

We can now estimate the average prestress as

$$\sigma_0^{crack} \approx \frac{1}{2}\overline{\Delta\sigma} + \frac{\Delta E_W^{crack}}{\overline{D}S} \qquad 8.89$$

At this point, we can use 8.84 to estimate $\Delta E_W^{crack} = E_f^{crack} + E_G^{crack} + E_R$. Although we need to assume spatial smoothness of stress and slip to estimate friction and fracture energy, radiated seismic energy is directly inferred from seismograms; assuming spatial smoothness is not necessary for this calculation. Although friction energy cannot be inferred from seismograms, we can nevertheless put upper limits on its size. The same goes for fracture energy.

FRICTION

Sliding friction during an earthquake is often called dynamic friction, σ_f . Friction is a dissipative process; it cannot add energy into the mechanical system. In the case of a uniform crack, $\sigma_f = \sigma_0 - \Delta\sigma$. Laboratory measurements of sliding friction for materials found in fault zones typically indicate that $\sigma_f \approx \mu_f \sigma_n$, where μ_f is the coefficient of friction and σ_n is the compressive stress normal to the fault. Most fault materials have measured coefficients of friction between 0.6 and 0.8. The normal stress on the fault is typically thought to be comparable to the pressure from the weight of the rocks above the earthquakes; this is referred to as lithostatic pressure and it is estimated to be $\sigma_n \approx \rho g z$ where z is depth beneath the Earth's surface. Assuming, the average density of crustal rocks (about 2,700 kg per cubic meter), this can be approximated as $\sigma_n \approx z \times \dfrac{30\text{MPa}}{\text{km}}$. Lithostatic pressure contrasts with hydrostatic pressure (the pressure of water in an open well) which is $\sigma_H \approx z \times \dfrac{10\text{MPa}}{\text{km}}$. If a material is porous and saturated with water, then the effective pressure is $p_{eff} = \sigma_n - \sigma_H$. The bottom line is that laboratory friction measurements indicate that the friction is on the order of $\sigma_f \approx z \times \dfrac{20\text{MPa}}{\text{km}}$. That is, at 10 km, we expect that frictional stress is on the order of 200 MPa. This expected frictional stress is much larger than the average measured stress drop of earthquakes (3 MPa).
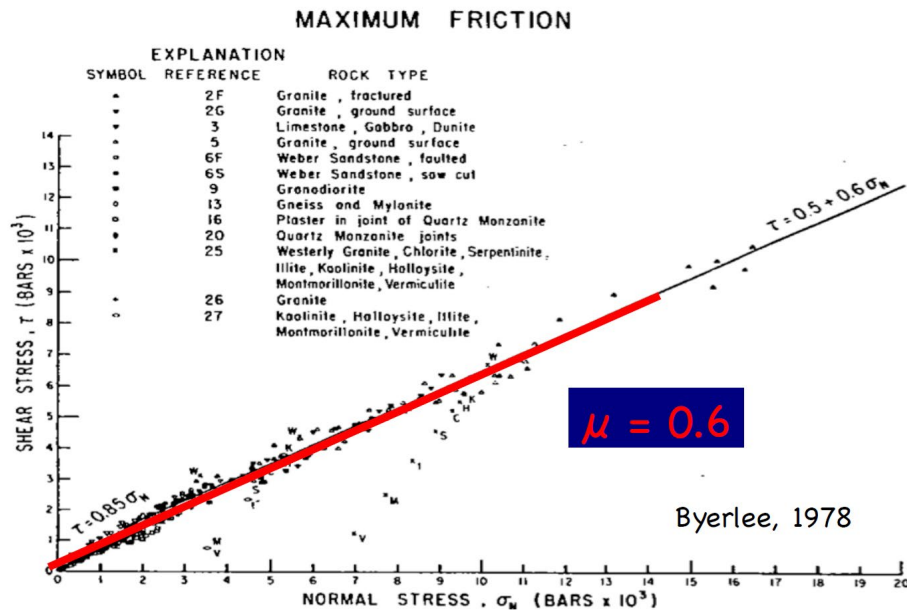
Figure 8.25. Shear stress of frictional yielding of a large range of experiments on different rocks and different normal stresses as reported by Byerlee. Although there are some notable exceptions (especially clay materials), most rocks have an apparent coefficient of friction of about 0.6. Note that the stress is given in units of kilobars (= 100MPa).

Figure 8.25 is a famous figure that shows the apparent coefficient of friction for a wide range of Crustal rocks subjected to a large range of confining stresses. This observation leads to one of the most fundamental unanswered questions in geophysics. What is the average shear stress in the seismogenic crust? Is it near the laboratory observations ($\sim$ 200 MPa), or is it comparable to the stress drop in earthquakes ($\sim$ 3 to 10 MPa)?

One clear observation that crustal rocks can sustain high shear stress comes from the fact that there are deep mines (up to 3 km in South Africa). At 3 km, the overburden stress is about 100 MPa. On the other hand, the interior surface of a mine is traction free. The calculation of stress around a mine is a complex problem, but it is clear that the shear stresses associated with deep mines are at least 50 MPa, which is consistent with Figure 8.21. Deep mines cannot remain open unless there are high shear stresses in the vicinity of the mine faces. Tragically, sometimes these faces fail violently in a phenomenon known as rock bursts. Rock bursts kill many miners in the deep gold mines of South Africa.

Figure 8.26 shows a popular view of the strength of the Earth's crust. This cartoon is often called the "Christmas tree." The maximum shear stress before frictional sliding is often called the "Byerly limit;" when you hear this, it is intended to refer to the Christmas tree strength model. I have often heard earth scientists use the term that the shallow crust is both brittle and strong, but the lower crust is ductile and weak. If you remember our discussions about the properties of buildings (Chapter 6), you should be confused that brittle is used together with the word strong. Remember that in structural engineering, the goal

is to avoid brittleness since local flaws can cause global collapse. Most structural engineers try to make their buildings ductile.
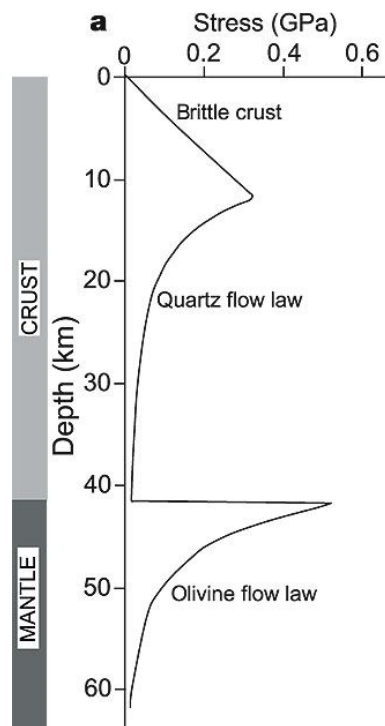

Figure 8.26 shows a simplified picture of the "strength" of the crust and is often called the "Christmas tree." The linear increase in the top 10 km assumes a constant coefficient of friction taken together with the assumption that pressure increases linearly with depth. The decreased stress in the lower crust is assumed to be the result of ductile yielding of rocks at high temperatures present in the lower crust.

Interestingly, mechanical engineers who work on fracture mechanics often describe the strength of brittle materials with either the fracture energy, or alternatively the stress intensity factor, $K_{III} = \sqrt{2\mu G}$. Note that fracture energy has units of $\dfrac{force \times length}{rupture\_area}$, whereas yield stress has units of $\dfrac{force}{rupture\_area}$. That is, these two different descriptions of strength differ in units by a length scale. Interestingly, brittle materials typically fail on thin surfaces (length scale of L^2), whereas failure in ductile materials is typically distributed in a volume (length scale of L^3). In many real-world cases, faults are complex at all length scales (apparently fractal). That is we can anticipate that the complexity of fault surfaces implies that we will have to resort to descriptions of strength that are somewhere between brittle and ductile. These laws will have units of stress times length to a fractal power. If the power is zero, L^0, then the material will have a simple yield stress (like a plastic material). If the power is 1, L^1, then the material is brittle and its strength is described with a fracture energy. That is, we

can already anticipate that there will be length dependencies to strength and that spatial complexity will be part of the story.

Friction Laws

Earthquakes appear to be due to shear sliding along fault planes; that is, inelastic strains associated with earthquakes occur in very narrow regions, which are often simplified as a planar interface. Although the inelastic strains almost certainly occur over a finite width, this width may be very small compared to other rupture dimensions. In order to construct simple models of fault rupture dynamics, we often assume that there is indeed slipping on a plane and that the shear stress on the plane is governed by a friction law. If the shear traction on a fault plane is less than the sliding friction, then the boundary is assumed to be a welded contact for which all displacements are assumed to be continuous across the fault plane. As long as the amplitude of stress is less than the friction law, the stresses on the fault plane are described by the standard laws of continuum mechanics (Chapter 3). However, the stresses cannot exceed the friction law; once the deformations are large enough, then the displacements on the fault are no longer continuous across the fault and the shear tractions on the fault are assumed to be equal to a friction law. These friction traction vectors are always assumed parallel to fault slip velocity, but in the opposite direction. The simplest possible friction law is just an instantaneous jump in friction from μ_1 to μ_0 . In this case, if the shear traction rises to $\mu_0 \sigma_n$, then the shear traction instantaneously drops to $\mu_1 \sigma_n$ and there is an instantaneous stress drop of $\Delta\sigma = \sigma_n \left(\mu_0 - \mu_1 \right)$, that is, assuming that the normal stress is constant. This was the assumption that Brune made in his 1970 paper.

Most observations of friction show a linear increase in frictional traction with fault-normal compressive stress. Examinations of machined surfaces typically show that contact surfaces are never truly planar; instead they all have geometric roughness (typically fractal). When two rough surfaces are placed into contact, then the actual contacts occur at only a few places (if the material is rigid, then only three contact points are enough to form a statically stable configuration). Of course, the normal force on the materials would be concentrated at these contact points. The result is that the contact points elasto-plastically deform into contact patches, that are typically called *asperities* (see Figure 8.23). The frictional shear tractions are also concentrated at these asperities. One common view is that the static coefficient of friction is simply describing how the total area of asperities changes as the normal stress is increased.

This view that friction laws are essentially describing the properties of surface asperities seems consistent with the observation that fault normal compressive stress is best measured by also including the effect of fluid pressures. That is, any interstitial fluids in the rough fault surface serve to push the surfaces apart. Obviously, this argument works well if the area of asperities is small compared to the overall area of the surface. If there are almost no void spaces, then the effect of fluid pressure is mitigated.
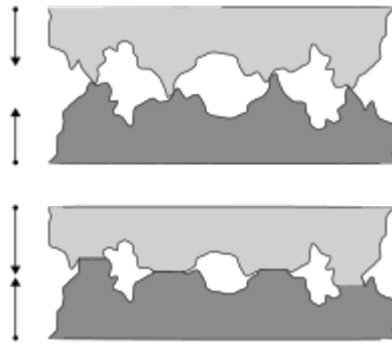
Figure 8.27. The contact area of geometric asperities are thought to be the key aspect of standard Byerly friction, The contact area increases with the normal stress
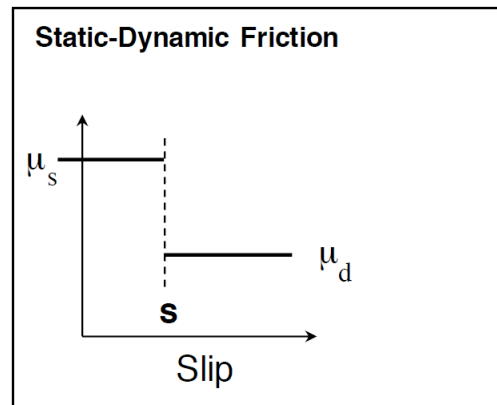

Figure 8.28. Instantaneous stress drop friction law where the friction drops instantaneously from a static coefficient of friction to a dynamic coefficient of friction. This law is unphysical because the properties of the material change without doing any work.

Although this "instantaneous stress drop" friction law is convenient for its simplicity, it has a fundamental and fatal flaw. That is, the material changes its physical properties without doing any work to the material. In the case of an instantaneous stress drop law, there is no fracture energy, which means that the material is infinitely brittle. A far more realistic situation is one in which the properties of a material change as a result of work that is done on the material. It may be that this work is small compared to other energies, but there must always be some work associated with changing the properties of a material. Realistic friction laws always involve parameters that describe the "state" of the material. In particular, as inelastic yielding occurs, the material's yield stress evolves as well. If we assert a simple friction law that changes from an initial friction to a final friction, it is necessary that there is some evolution of this change; it cannot be instantaneous.

Simple Slip-Weakening Friction

Slip-weakening friction as shown in Figure 8.29 is commonly used in simulations of earthquake rupture. It has the advantages that there is work associated with change in the frictional state of the material. That is, the work associated with the transition from static friction to sliding friction is just the area of the back-hatched region shown in 8.29. Slip-weakening friction is described by

$$\tau_f \leq \sigma_n \begin{cases} \mu_S - (\mu_S - \mu_d)\dfrac{D}{D_0} & 0 < D < D_0 \\ \\ \mu_d & D > D_0 \end{cases} \qquad 8.90$$

Where D_0 is the **slip-weakening distance**. The fracture energy (per unit of rupture area), G_0 , is given by

$$G_0 = \frac{1}{2}\sigma_n D_0 (\mu_s - \mu_d) \qquad 8.91$$

and the frictional heat energy (per unit of rupture area), E_F , is

$$E_F = \sigma_n \mu_d D \qquad 8.92$$
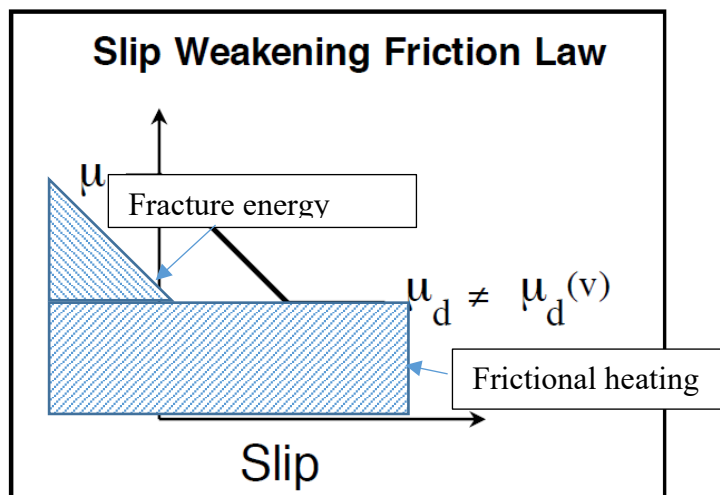
Figure 8.29. Slip-weakening friction law is common in rupture dynamics.   The lower hashed area represents the total work done by frictional heating $\sigma_o D$, while the upper back hashed area represents the work done to drop the friction and is commonly interpreted as fracture energy.

As long as $D > D_0$, slip-weakening produces a simple stress drop. However, this law also includes a length scale that is associated with changing the frictional state from static to dynamic. In fact, any friction law that changes the internal energy of a material has included (either explicitly or implicitly) a characteristic length in the friction model. Rather interestingly, **Aki's conjecture of scale-free rupture**

dynamics is violated by any friction law with a scale-independent fracture energy.

Notice that slip-weakening friction specifies the maximum shear that can occur before sliding, $\mu_s \sigma_n$. Sometimes this number is called the strength of the fault since it is the maximum shear stress that can be achieved locally. While this parameter is relevant to the conditions of crack nucleation, it turns out that it has virtually no effect on the dynamics of propagating failures. Instead, propagating failures are more related to the work necessary to extend a rupture, which is the fracture energy. Fracture energy is the product of peak stress and slip weakening distance, and neither quantity can be independently determined from dynamic modeling of rupture (Gutierri and Spudich).

A major problem with slip-weakening friction centers on the problem of rupture nucleation. That is, in order to have a seismic instability, the rate at which potential strain energy is extracted by the rupture must exceed the rate at which energy is dissipated by friction and fracture on the expanding rupture surface. We can approximate the change in potential energy due to faulting that is available for radiation, $\Delta E_{W-f}(t)$, as the change in potential energy minus the frictional energy as

$$\Delta E_{W-f}(t) \approx S(t)\,\overline{D}(t)\,\sigma_{eff}$$

$$\approx S^{3/2} \frac{\Delta\sigma}{C\mu}\sigma_{eff} \qquad\qquad 8.93$$

Where S is the rupture area, $\overline{D}$ is the slip averaged over the rupture area (scales with L, see 8.69), and σ_{eff} is the difference between the prestress and the dynamic friction (see 8.7). That is, it is the effective stress that is available for dynamics.

I have written this as an approximate relationship because I am assuming that the effective stress is constant in space. Later we will see that the prestress, the slip, and the dynamic friction are spatially heterogeneous. 8.93 needs to be modified if it correctly considers this heterogeneity.

Now the energy that is available for radiation E_R (kinetic energy) is just

$$E_R = \Delta E_{W-f} - E_G$$

$$\approx S^{3/2}\frac{\Delta\sigma}{C\mu}\sigma_{eff} - SG_0 \qquad\qquad 8.94$$

$$= S\left(\frac{\Delta\sigma}{C\mu}\sigma_{eff}\sqrt{S} - G_0\right)$$

where G_0 is the fracture energy per unit rupture area. This means that in order **to have spontaneous sliding** , there must be more change in potential energy than is required for fracture and friction, Or,

$$\sqrt{S} > \frac{G_0 C\mu}{\Delta\sigma\sigma_{eff}} \qquad\qquad 8.95$$

That is, slip weakening friction implies that if the prestress is homogeneous, then there is a minimum dimension for earthquakes. If we use the laboratory measures of fracture energy (say 200 J/m^2), and effective stress of 2 MPa (estimated from seismic energy radiation), a typical stress drop of 2.7 MPa, then the minimum rupture length for an earthquake is about a meter, which is very small. However, if we use fracture energies that are required to stop a large-dimension crack in a uniform stress field (see 8.35) then we end up with fracture energies of tens of MJ/m^2 . In this case, it is necessary to use high prestress in the nucleation region just to get things started. Currently, many researchers insert patches of high pre-stress in order to nucleate instabilities. While inserting high-stress patches is easy on a computer model, it requires a religious miracle to realize in the real world.

Unfortunately, many researchers report the fracture energies that are employed in their dynamic modeling and interpret this energy as having physical significance. They rarely mention that using small fracture energies causes extreme numerical instability in their dynamic calculations. That is, in many cases the reported fracture energy is the energy that is required to produce stable finite-element models (a function of the grid spacing and time step). Small fracture energies produce singular stresses at the crack front. Discrete element codes cannot handle these singular stresses, and the modelers increase the fracture energy in order to soften the singularity at the crack tip. In addition, the use of small fracture energy means that this parameter has almost no impact on the details of the fracture dynamics. Basically, fracture energy is used to stop fractures. However, it is impossible to stop long cracks without very large (stupidly large) fracture energies.

Another consequence of using fracture energy together with crack-like dynamics is the fact that discrete numerical simulations are not only poorly resolved, they also typically jump to super-shear rupture velocities as the rupture dimension grows. In discrete models (e.g. finite elements) with low fracture energy, the expanding crack fronts evolves into such large stresses that elements several grids in front of the crack can begin to slip even before the crack front arrives. That is, the rupture jumps ahead of the crack front. This is super-shear rupture velocity and one it starts, it can propagate stably. Although there have been reports of evidence for sup-shear rupture velocities in some earthquakes, most well recorded earthquakes appear to have sub-shear rupture velocities. Given the behavior I just described, the real mystery is to explain why all significant earthquakes don't have super-shear rupture velocities. I will shortly show that a better way to control the rupture dynamics is with the prestress.

Rate and State Friction

Although instantaneous weakening friction and slip-weakening friction have been commonly used in earthquake dynamics because they are so simple, they are not actually based on experimental observations of frictional sliding. Frictional sliding is actually a specialized type of localized inelastic shear straining. In general, the inelastic shear strain rate depends on the applied shear stress and it also depends on the past history of inelastic strain (the "state" of the

interface) and it can also depend on the current rate of sliding. Friction laws that have these two dependencies belong to a class of laws called **rate and state friction laws.** The use of variables that characterize the state of a deforming material are quite common in material science; for example, the stress-strain relationship for reinforced concrete is very complex. In particular, it is important to know if the concrete has been cracked in previous straining. In that sense, the amount of fracturing is a state variable for reinforced concrete.

In the Earth Sciences, the term, "Rate and State friction" has been used almost exclusively to mean a particular friction law that was developed to describe the temporal behavior of a number of low-slip-rate (less than 1 mm/s) frictional sliding experiments. I personally prefer to use the term "Dieterich-Ruina friction" when referring to friction laws of the type described below. Since rate and state dependent rheologies are common in mechanics, it's easy to be confused when using the phrase rate-state friction.

In a foundational set of friction experiments, James Dieterich (1979, Modeling of rock friction 1. Experimental results and constitutive equations, JGR, 84, 2161-2168) observed the time variation of slip along the interface of machined blocks of granite that were subjected to time-dependent shear stress and constant fault-normal compression. Dieterich discovered that if a block was forced to slide at a steady slip rate, then an instantaneous step in the shear traction caused a change in the slip rate and that this change in slip rate was not instantaneous, but instead evolved over some time scale (or alternatively a slip scale). Later, Andrew Ruina (1983, Slip instability and state variable friction laws, JGR, 88, 10,359-10370) introduced a state variable that could predict this behavior.

The Dieterich-Ruina friction law is briefly summarized in Figures 8.26 and 8.27, which I copied from a presentation by Chris Marone at Penn St. This friction law provided a new framework for studying fault dynamics. Most importantly, the definition of several constants $\left(A,B,V_0,D_c \right)$ accurately mimics the experimentally observed slip rate as a function of time for a wide variety of materials. This friction law is often used as the basis of numerical modeling of fault slippage in a wide range of problems. Notice that the Dieterich-Ruina friction has a state variable, $\Theta(t)$, which is described by a 1$^{\text{st}}$ order differential equation in either time (or alternatively slip). $\Theta(t)$ has units of time and it controls how the friction evolves after jumps in slip velocity (or alternatively, shear stress). The evolution after a jump evolves logarithmically (either time or slip) towards a new steady state. Clever experiments that allow one to observe the evolution of contact patches during slip have revealed that the fact that the asperities are constantly evolving on a slipping surface. Slow slip velocities allow dislocation creep to increase the asperity area (indentation creep), thereby increasing the friction. In this case, the friction decreases with increasing sliding velocity (but only after the effect of the state variable has decreased. On the other hand, increasing the slip velocity can cause higher temperatures at the asperities, which allows them to relax faster and thereby increasing the friction when the sliding rate is increased. These two effects are included in the Dieterich Ruina friction law as the constants, A and B. If A > B, then the Dieterich Ruina law predicts that increasing the sliding rate results in lower friction. This is

called rate-weakening friction and it produces dynamic instabilities in the sliding process. That is, if there is a perturbation that causes the sliding rate to increase, then this decreases the friction, which in turn, increases the sliding rate. This dynamic instability naturally results in earthquake nucleation.

On the other hand, if $A < B$, then faster sliding results in higher friction and is referred to as rate-strengthening friction. Perturbations to the sliding rate tend to damp out with time. I often see that faults that fail in earthquakes are interpreted as having $A > B$, whereas faults that experience creep are interpreted as having $B > A$.

It is worth noting however, that $B > A$ seems to imply that a fault creeps steadily. It is difficult to observe spatial and temporal variations in creep (there are no radiated waves and the creep is deep in the Crust). However, when creep has been observed in detail, it's my experience that it mainly occurs during creep events. These creep events seem to be much like other earthquakes, but with a greatly reduced rupture velocity (usually less than 1 km/hr).

There is a very large literature that discusses the attributes of Dieterich-Ruina friction and these class notes are far too brief to really do this subject justice. I recommend reading the review paper by Rice, Lapusta, and Ranjith (Rate and state dependent friction and the stability of sliding between elastically deformable solids, J. of the Mech. And Phys. Of Solids, 2001, 49, 1865-1898).

One important cautionary note concerns the range of laboratory conditions that were observed when creating this friction law. In particular, the experiments observed at slip rates of microns per s and observed characteristic slip distances, D_c, of microns. Furthermore, the changes in friction were small changes about average frictions of about 0.7. That is to say, the dynamic friction had similar values to traditional Byerlee friction, with the important addition of time evolution that provides a natural pathway to event nucleation. It's worth noting that when these experiments were sped up to earthquake slip velocities (> 1 m/s), then frictional melting always occurred. Once the material surfaces melts, these friction laws no longer apply. That is, while rate-weakening Dieterich-Ruina friction appears to be a reasonable model for the nucleation of earthquakes, it is inconsistent with observations of large earthquakes. Furthermore rate-strengthening friction provides a convenient model for creeping faults, but it fails to predict the common observation that creeping faults typically fail in slow earthquakes.

Rate (v) and State (θ) Friction Constitutive Laws

$$\mu(\theta, V) = \mu_o + a\, ln\left(\frac{V}{V_o}\right) + b\, ln\left(\frac{V_o\theta}{D_c}\right)$$

state variable, characterizes physical state of surface or shearing region

critical slip distance

reference velocity

reference value of base friction

$$\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_c}$$

Dieterich, aging law

$$\frac{d\theta}{dt} = -\frac{V\theta}{D_c} ln(\frac{V\theta}{D_c})$$

Ruina, slip law
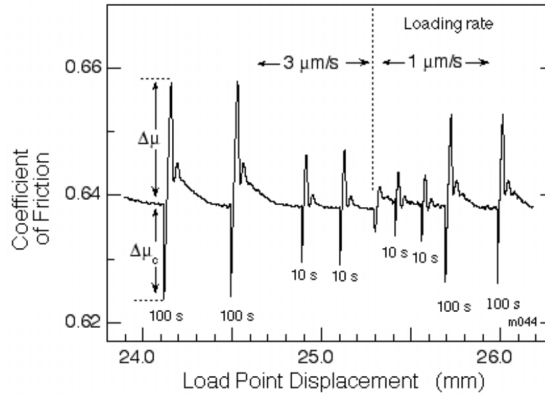
Figure 8.30

Rate (v) and State (θ) Friction Constitutive Laws

1) $\mu(\theta, V) = \mu_o + a\, ln\left(\frac{V}{V_o}\right) + b\, ln\left(\frac{V_o\theta}{D_c}\right)$

2) $\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_c}$

Convention is to use a, b for friction and A, B for Stress

$$\tau(\theta, v) = \tau_o + A\, ln\left(\frac{V}{V_o}\right) + B\, ln\left(\frac{V_o\theta}{D_c}\right)$$

$$A - B = \frac{\Delta\tau}{\Delta lnV}$$

Steady-state velocity strengthening if a-b > 0, velocity weakening if a-b < 0
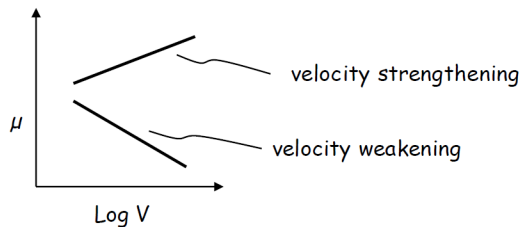

velocity strengthening

velocity weakening

Figure 8.31

Frictional Heating

Assuming either instantaneous weakening friction, slip-weakening friction, or even Dieterich-Ruina friction, we can estimate the average long-term rate of frictional heating (or power per unit rupture area) on a fault to be

$$P_F = \bar{D}\sigma_n \bar{\mu}_d \qquad\qquad 8.96$$

Where $\bar{D}$ is the average geologic slip rate of a fault and $\bar{\mu}_d$ is the average dynamic coefficient of friction. It is a simple calculation to estimate the heat production on the San Andreas fault assuming that the dynamic coefficient of friction is about 0.6 (i.e., Byerlee friction). Assuming that the average dynamic friction on the San Andreas is about 200 MPa (see Figure 8.22), and that the slip rate is about 3 cm/yr (or about 10^{-9} m/s) gives $P_F \approx 200$ mW/m^2 . This is actually a problematically large power; the average power of heat transferred through the Earth's continental crust (otherwise known as the heat flow) is 66 mW/m^2 (Davies and Davies, 2010). That is, the assumption of Byerlee friction implies that the average rate of frictional heat produced by the San Andreas fault is several times the average heat flow in the Crust and thus we would expect to observe elevated heat flow in the vicinity (within 10's of km) of the San Andreas fault. This elevated heat flow is not observed and this fundamental problem is often called the **Heat-flow Paradox** (see for example, Scholz, C. H., The Mechanics of earthquakes and Faulting, Cambridge University Press, Cambridge England, 1990).

Actually, the heat flow paradox is far more severe than is indicated by the calculation above. In particular, a variety of observations indicate that average slip rates during earthquakes are at least 2 m/s. This means that, during the slip event, if dynamic friction is about 200 MPa, then we expect frictional power of at least 400 MW/m^2 during the sliding. This is an extraordinarily large number. In fact, it's the energy of 200 sticks of dynamite per second applied to every square meter of the fault. Or, for comparison, the power of a nuclear power reactor is about 1 GW. The specific heat of granite is 790 J/^{o}Ckg, and the density is 2.7x10^{3}kg/m^3. Assuming that a temperature change of 1,000^{o}C would cause melting (rocks have a very low latent heat of melting), then the heat $E_{1mm\,melt}$ required to raise the temperature of a 1mm thick section of the fault by 1,000^o C is

$$E_{1mm\,melt} \approx 4 \times 10^8 \left(\frac{J}{s}\right) T_S$$

$$\approx 2.7 \times 10^3 \left(\frac{kg}{m^3}\right) \times \left(\frac{m}{10^3 mm}\right) \times 8.2 \times 10^2 \left(\frac{J}{^0 kg}\right) \times 10^{3\ 0} \qquad 8.97$$

$$\approx 2.2 \times 10^5 \left(\frac{J}{mm\text{-}m^2}\right)$$

Now the total frictional heat energy is

$$E_{Friction} \approx 4 \times 10^8 \left(\frac{J}{m^2 s}\right) T_S \qquad\qquad 8.98$$

Where $T_S \equiv$ duration of slip at a point . If $T_S = \frac{1}{2}s$, then the slip is about 1 m and there is enough frictional heat to raise a 2-m thick zone by 1,000°. Clearly, any coefficient of sliding friction larger than 0.05 would cause instantaneous melting of the fault-zone materials. This fact was originally pointed out by Harold Jefferies and a more detailed calculation was described by Richards (197?).

Although there are geological observations of melts observed in fault zones that have been exhumed by tectonic uplift followed by erosion of overburden materials (see for example, the Punchbowl Fault described by Chester and Chester, 199?), the total volume of melt is very small compared with the amounts expected from ordinary Byerlee friction. In fact, once a melt layer is developed, one expects the fault tractions to be controlled by the viscosity of the melt. Furthermore, we would expect to see that the final fault shear traction should drop to zero, which is often referred to as complete stress drop. A 200 MPa stress drop is a factor of one hundred larger than the average observed in major earthquakes. This means that the average slip should be a factor of 100 times larger than is typically observed.

Notice that the estimate of $200 \, \text{kJ}/\text{mm-m}^2$ to melt a mm-thick layer is an order of magnitude smaller than estimates of the far-field radiated energy from a earthquake with a 1-m slip, about $2D \, \text{MJ}/\text{m}_{\text{slip}}\text{-m}^2$ (see 8.42). That is to say that we should expect to find extensive frictional melting in the vicinity of active faults. The big mystery is, why they are not observed.

At this point, I will say that the paucity of melted rocks in exhumed faults implies that the total frictional heat per unit rupture is less than about 200 kJ/m². This implies very low dynamic friction for meter-sized slips. At this time, such a low dynamic friction has not been observed in a laboratory experiment. Which is to say that although it's clear that dynamic friction in large earthquakes is very low, we don't know why.

Strong Rate-Weakening Friction

In my 1990 slip-pulse paper, I argued that strong rate-weakening friction could be the key to understanding the origin of slip pulses. I argued that it could also help to explain the low heat observed from tectonic deformation.

One simple realization of a rate weakening friction is to hypothesize that

$$\mu = \frac{\mu_0}{1 + \left(\dot{D}/\dot{D}_0 \right)^{\beta}}$$

8.99

(see Cochard and Madariaga, 199?). In this friction law, the friction has a power law dependence on slip velocity. There is no inherent to distance in this friction

law, and hence this friction law cannot provide any inherent length scale to rupture problems that are described by this law. While none of the friction laws discussed to this point are capable of generating slip pulses, this friction law can. In order to estimate how much less heat can be produced by this friction law, we integrate the frictional force over the slip in a slip pulse, or

$$E_F = \int_0^{T_0} T_F(t)D(t)\,dt = \int_0^{T_p} \left(\frac{\mu_0}{1+\left(\dot{D}/\dot{D}_0\right)^\beta} \sigma_n D_p(t) \right) dt \qquad 8.100$$

Now assuming that the slip history of a pulse can be approximated as a truncated growing crack, then

$$D_p(\tau_p) \approx \begin{cases} D_0\sqrt{\tau_p} & 0 < \tau_p < 1 \\ D_0 & \tau_p > 1 \end{cases} \qquad 8.101$$

where $\tau_p \equiv t/T_p$. And then $dt = T_p d\tau_p$, and

$$\dot{D}(\tau_p) \approx \frac{1}{2}\frac{D_0}{\sqrt{\tau_p}} \qquad 8.102$$

Substituting 8.101 and 8.102 into 8.100 gives

$$E_F = \sigma_n \mu_0 D_0 \int_0^1 \frac{1}{1+\left(\dfrac{\dot{D}(\tau_p)}{\dot{D}_0}\right)^\beta} \sqrt{\tau_p}\,d\tau_p$$

$$= \sigma_n \mu_0 D_0 \int_0^1 \frac{1}{1+\left(\dfrac{D_0}{\dot{D}_0}\dfrac{1}{2\sqrt{\tau_p}}\right)^\beta} \sqrt{\tau_p}\,d\tau_p \qquad 8.103$$

Or

$$E_F = \sigma_n \mu_0 D_0 F_{rw}(\beta) \qquad 8.104$$

Where F_{rw} is a factor that includes the effect of power-law rate-weakening and

$$F_{rw}(\beta) = \int_0^1 \frac{4\tau_p}{1+\left(\dfrac{1}{\sqrt{\tau_p}}\right)^\beta}\,d\tau_p \qquad 8.105$$

Unfortunately, I have not been able to find an analytic solution to this integral. However, this definite integral can be numerically evaluated and the answer is shown in table 8.2.

β	F_{RW}
0.1	0.98
0.5	0.94
1	0.88
2	0.77
4	0.61
8	0.43
16	0.27
32	0.15
64	0.081
128	0.042

Table 8.2

Table 8.2 seems to imply that the use of the rate-weakening law still results in substantial frictional heating. That is, unless the law assumes very strong rate weakening (i.e., $\beta > 20$). The paucity of melt observed in exhumed faults is strong evidence that dynamic friction drops very rapidly behind the rupture front. Perhaps it is a sudden transition from high static friction $\left(\mu_{static} \sim 0.6\right)$ to very low dynamic friction $\left(\mu_{dynamic} < 0.05\right)$. At the time of these notes, I am unaware of experimental observations of such low dynamic friction.

Dynamic Friction in Landslides sudden transition to low dynamic friction

One of the most direct ways to observe a rapid transition from high static friction to very low dynamic friction is to study the dynamics of landslides. For example, Fig. 8-32 is a photograph of Mt. St. Helens just seconds prior to its eruption in 1980. Mt. St. Helens is a typical composite volcano. These volcanoes are conical structures that form by the accumulation of pumice, ash, and volcanic bombs that eject from the top of the volcano. In essence, a composite volcano is simply a large pile of unconsolidated material that is sitting at its angle of repose. The angle of repose is the steepest slope that a pile of granular material can obtain in static equilibrium. The angle of repose is determined by the coefficient of friction between the granular materials. In the case of Mt. St. Helens, the volcanic cone was in static equilibrium with a slope of about $\theta \approx 29^o$. The inter-granular coefficient of friction of a marginally stable cone of material is $\mu_f \approx \tan\theta$, or in the case of Mt. St. Helens, $\mu_f \approx 0.54$. Steady inflation of the magma beneath the volcano occurred for a period of several months prior to the eruption. This inflation (about a 100-m bulge) over steepened the slope. About 10 seconds prior to the eruption, a M 5.1 earthquake shook the over-steepened slope and an enormous landslide was triggered. Figure 8-33 is a photograph taken early in the landslide. A remarkable time lapse of photographs provides data to estimate the velocity of the sliding mass .
https://www.bing.com/videos/search?q=mt+st+helens+time+lapse+video&qpvt=mt+st+helens+time+lapse+video&view=detail&mid=58ACA9463AB2BDCC740F58ACA9463AB2BDCC740F&&FORM=VRDGAR

In the time lapse, it can be seen that approximately 3 km^3 of material slid downslope with a velocity of about 200 km/hr. The base of the slide appears to be on a steep slope of about $\beta \approx 50^o$. The acceleration of the landslip can be approximated as $\ddot{u} \approx g\left(1-\mu_{dynamic}\right)\sin\beta$. The time lapse indicates that the friction dropped almost instantaneously to near zero and that the landslide mass accelerated as if frictionless on the steep sliding plane. As the mass slid downslope, the pressure from the weight of the landslide mass was decreased in the magma chamber (very high viscosity alkaline magma) beneath the mountain. This caused rapid outgassing of volatiles that were dissolved in the previously confined magma. This rapid outgassing caused the lateral blast phase of the eruption.
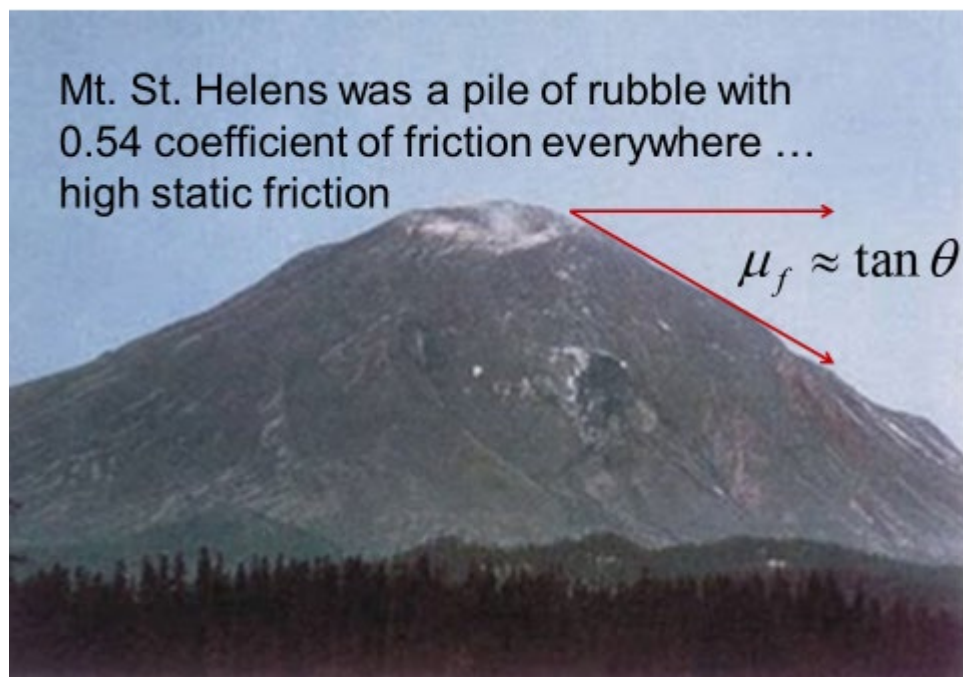

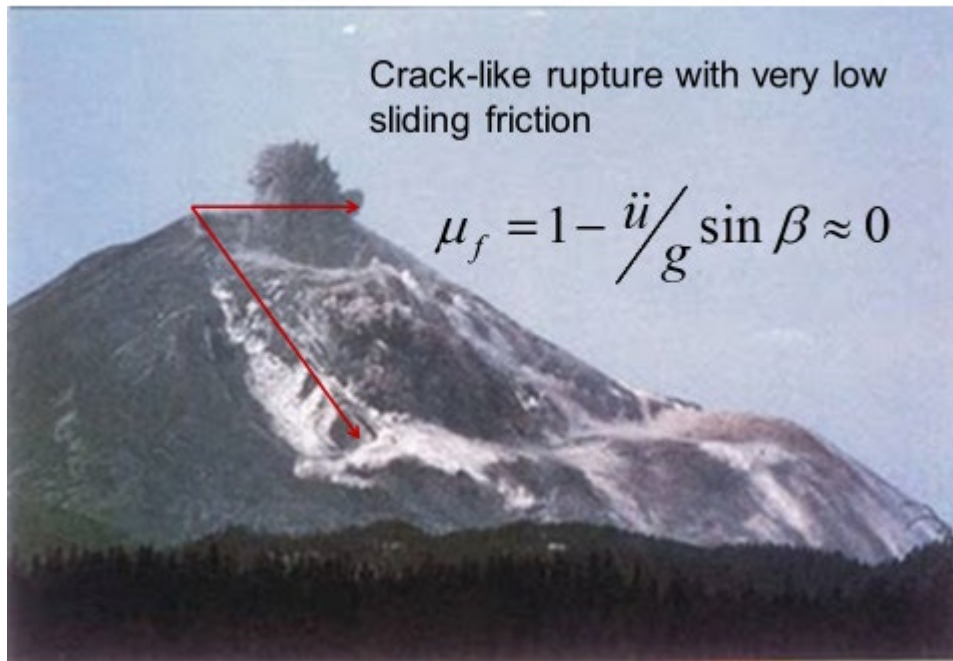Figure 8.32. Mt. St. Helens volcano prior to its eruption in 1980.

Figure 8.33. Mt. St. Helens during the massive landslide that depressurized the magma chamber.

Landslides are very similar to earthquakes; they are dynamic shear sliding on a slip plane. However, in earthquakes, the rupture is confined to a portion of the fault plane, while in landslides, there are no ends of the rupture plane. In essence, landslides are crack-like; the entire fault plane slips simultaneously. The downslope slips can become very large since gravity provides the driving mechanism. In an earthquake the fault-plane shear stress decreases as slip increases. In landslides, the stress does not decrease until the mass is at the bottom of the hill.

The most important point of Figure 8.33 is that the static friction at Mt. St. Helens was at least 0.54 prior to the M 5.1 earthquake. A failure plane with very low dynamic friction (less than 0.1) formed very rapidly during the massive landslide. If the rupture velocity was similar to earthquakes (V_R about 2 ½ km/s), then this plane formed in less than 1 s.

One simple way to parameterize the dynamic friction in a landslide is to estimate the ratio of the vertical displacement, H, to the horizontal displacement, X, of the center of mass. This is a measure of the average slope of the landslide, and the effective coefficient of dynamic friction is just H/X . Estimates of dynamic friction in landslides vary from almost 1.0 to as low as 0.06. There seems to be a clear indication that landslides with larger total volumes tend to have smaller dynamic frictions. $\mu_{dynamic} \approx H/X \approx 1/\sqrt[8]{1+V}$ (where V is landslide volume in m^3) approximately fits widely scattered observations (Johnson, Campbell, and Melosh, 2016, The reduction of friction in long runout landslides as an emergent phenomenon, JGR, 121, 881-889, doi/10.1002/2015JF003751).

Force Chain Networks

I will now briefly describe the behavior of granular particle systems. This will help to introduce the concepts of self-organization, emergent systems, and fluidization.

Sand is an example of a material that has a rich variety of behaviors. It is comprised of simple grains, each of which is described by the dynamics of a rigid body. However, a collection of interacting sand grains can display behavior that is difficult to anticipate based solely on the rigid body dynamics of each particle. For example, sand in an hourglass can flow like a viscous liquid. Sand can also interact with wind to form sand dunes that evolve and appear to slowly propagate through space. Sand can also compact into a closely packed material that is well suited to form the foundation of structures (condensed matter physicists call this a jammed granular material).

In the case of sand foundations, it's important to understand the conditions that cause inelastic yielding of the sand. One approach to collections of sand is to consider that the particles are so small and numerous that they can be simulated using a continuum (i.e., there is no smallest scale). Numerous tests of yielding sand has demonstrated that it can be described as an elasto-plastic continuum. In the case of sand, the plastic yield shear stress is dependent on the confining pressure of the sand. This is usually called Drucker-Prager plasticity. In contrast, steel is described with Von Mises plasticity, where the yield stress is independent of confining pressure.

The idea that there is a maximum shear stress that can be supported by sand seems intuitively reasonable, and you might believe that that is all you need to understand about the deformation of an assemblage of sand particles. However, there has been new research that is revealing surprising mechanics behind the simple statement that a sand is elasto-plastic. In a sense, it should not be unexpected that there is something else at play; sand is not elastoplastic at the grain level (it's elastic-brittle). It's only a large assemblage of sand particles that behave plastically.

Experiments that reveal the mechanics of interacting sand grains show that most of the load carried by sand is actually borne by a relatively small percentage of the sand grains. Furthermore, even if the external tractions may produce bulk shear of the sand assemblage, the actual loads are primarily axial stresses directed along collections of grains. These are called **force chains**. The details of how and when individual chains appear are very sensitive to particular shapes and orientations of individual grains. Although these systems were first studied using physical experiments, recent advances in computational dynamics have demonstrated that these types of systems can be numerically simulated (for example, see https://resolver.caltech.edu/CaltechTHESIS:12042014-104112714).

Fig. 8.34 is copied from the website of Robert Behrenger (Duke Univ. Physics Dept., deceased). The force chains carry axial forces that resist shearing force. The axial forces cause shear in the photosensitive beads. A video from this

experiment shows how finite motions result in eventual failure of individual force chains. After failure, new force chains rapidly develop and the ensemble behavior of the system appears as if it's a plastic material.

The force chains shown in Fig. 8-34 are a complex network that forms almost instantaneously. This network is an example of a **self-organizing system**. The formation of the force chains is natural and fundamental to understanding the mechanics of sand. Even though prediction of the details of the network of force chains is very difficult, you can be certain that appropriate force chains will occur.

Although the sand system deforms in a sequence of small jumps, the effect of individual chains is small enough that a system of many particles behaves similarly to traditional sliding friction. The system shown in Fig. 8.34 is an example of a **jammed granular system**. If the confining pressure of the system is removed then the particles can move freely and the same sand grains can show deformation that is similar to a viscous fluid, which is typically referred to as a **granular fluid**. Granular fluids are important in many industrial processes since this is a way to mix and transport granular materials. The transition between jammed and fluidized states can occur very quickly.

Fig. 8-35 is a photo taken after a massive landslide that was triggered by the M 7 ½ 1959 Hebgen Lake Earthquake. A substantial fraction of a 2.3 km high mountain slid into the Madison River Valley at an estimated speed of 150 km/hr. The momentum of the slide was sufficient to cause the slide to climb several hundred m above the opposite valley wall. As soon as the slide's velocity dropped to zero as it climbed the valley wall, it froze into place. This seems to be direct evidence of sudden transitions in the apparent friction when 1) it was initially stable on the mountain slope, 2) it suddenly transitioned to very low dynamic friction allowing the slide to accelerate downhill at high speed, and 3) suddenly transitioned back to high friction when the sliding velocity dropped to zero.
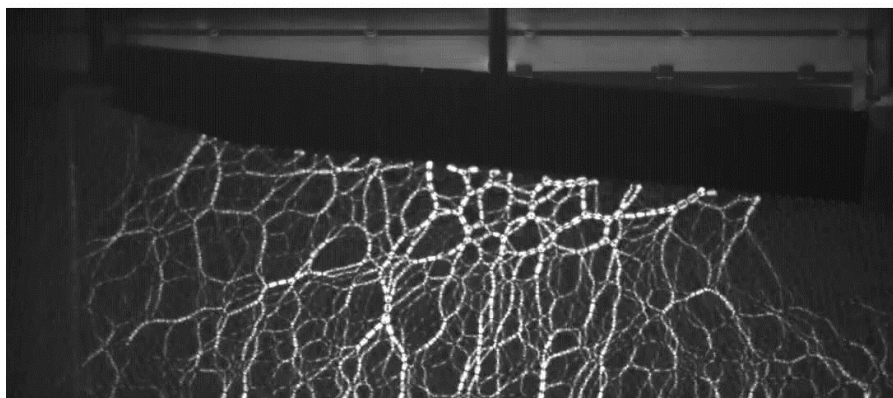

Fig. 8-34. Photograph of force chains that result from the shearing of an assemblage of photo-sensitive plastic beads that are confined to a rectangular bucket. The black wedge above is moving to the left. The polarized light is transmitted by beads experiencing shear strain that is the result of uniaxial stress carried by the force chains (from Robert Behrenger's web site).

Figure 8.35. Landslide across the Madison River in Montana triggered by the 1959 M 7 ½ Hebgen Lake earthquake. Note that the slide had sufficient momentum to climb the the opposite valley wall. I seems that the friction transitioned back to high static friction when the sliding velocity dropped to zero.

Some of the most dramatic examples of large landslides are from the Hawaiian Islands. The dotted lines in Figure 8.36 map out the areal extent of landslide deposits that originated from the slopes of Hawaiian volcanos (Moore, J., D. Clague, R. Holcomb, P. Lipman, W. Normark, and M. Torresan, 1989, Prodigious submarine landslides on the Hawaiian Ridge: Journal of Geophysical Research, v. 94, p. 17,465–17,484). Several of these enormous landslides appear to have run out more than 100 km from their source region. Furthermore, there is evidence that giant tsunamis (run-up heights exceeding several hundred meters) carried beach rocks far up the flanks of existing volcanos. The most recent of these events has been dated at about 100,000 ybp (Megatsunami deposits on Kohala volcano, Hawaii, from flank collapse of Mauna Loa, McMurty, and others, 2004, Geology, 32, 741-744, doi: 10.1130/G20642.1).

Figure 8.37 shows a map of the head scarp of one of these landslide structures (P Lipman, T. Sisson, and J. Kimura - 2-4. Hilina slump area, 2001, Public Domain, https://commons.wikimedia.org/w/index.php?curid=69511814). This scarp is called the Hilina slump (or alternatively, the Hilina fault). Figure 8.38 is a photograph of the Hilina escarpment. Obviously, the landslide is currently in equilibrium, which means that the friction in the landslide exceeds the shear forces that result from the gravitational weight of the slide.
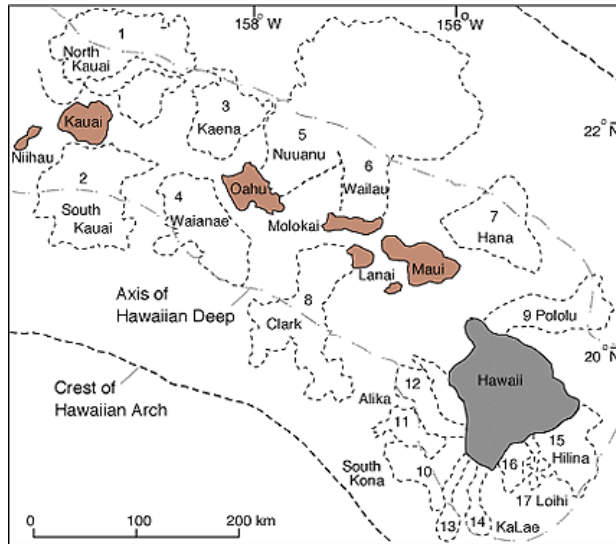
Fig. 8.36. Spatial extent of debris fields created during pre-historic collapse of the flanks of Hawaiian volcanoes. These enormous landslides sometimes ran more than 100 km along the ocean bottom. (Moore and others, 1989)
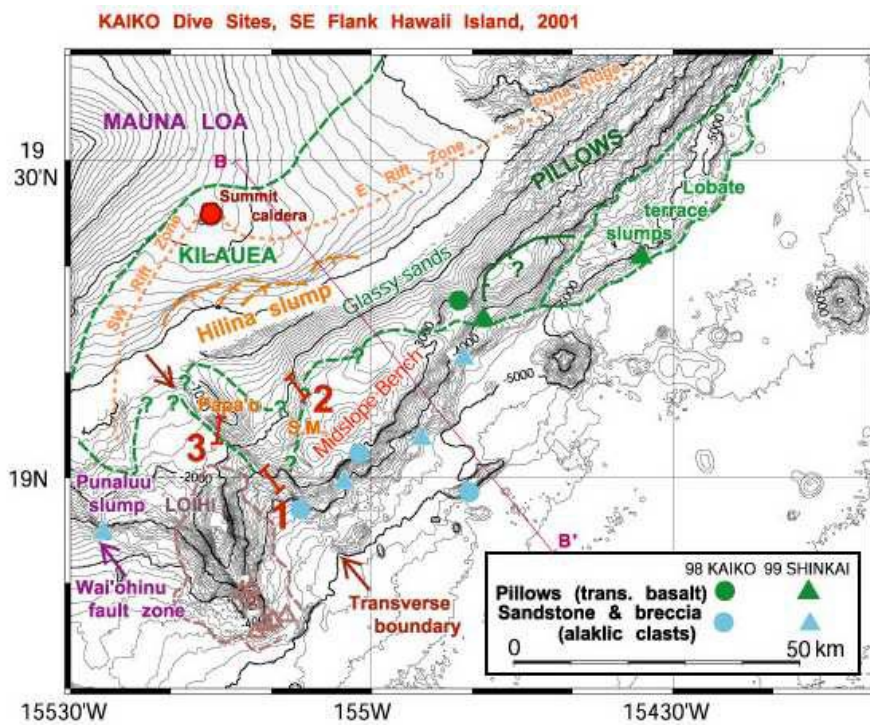

Figure 8.37 Map showing the extent of the Hilina slump which appears to be an incipient flank collapse structure.

Figure 8.38. Photograph of the escarpment of the Hilina fault (landslide).

Figure 8.39 shows a schematic cross section of the Hilina slump. The eruption of lavas built the broad swell of the Island of Hawaii. The old ocean floor sits at the base of volcanic deposits, which are more than 10 km thick in the center of the volcano. The weight of the volcanics have depressed the ocean floor so that the interface between the volcano and the ocean floor dips shallowly towards the volcano. This means that the toe of the landslide actually moves upwards due to gravitational collapse of the head of the landslide.
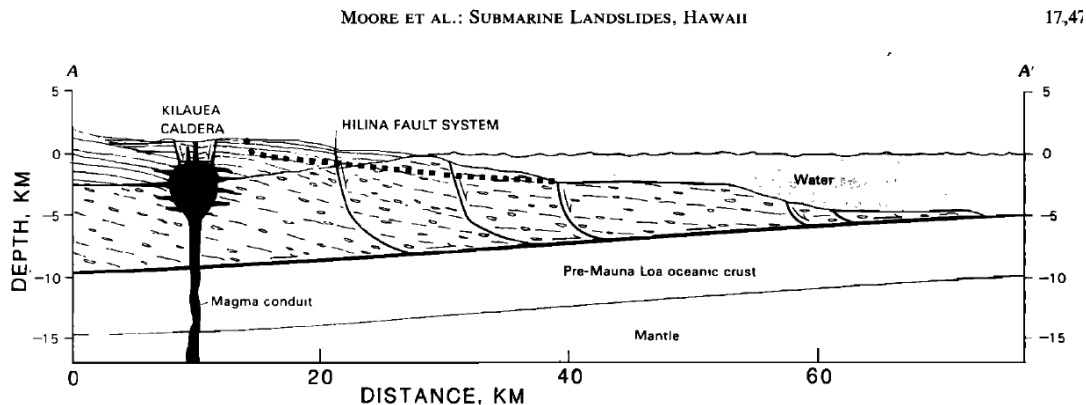
Fig. 11.   Section through Hilina slump with no vertical exaggeration modified from *Lipman et al.* [1985]. Location of section shown in Figure 10. Dashed pattern forming base of volcano is subaerially erupted clastic rocks including phreatic explosion ash, littoral cone ash, flow foot breccias, minor pillowed flows, and pillow breccias, landslide debris, and turbidity flow deposits. Most of this material was probably erupted from Mauna Loa except for the uppermost few kilometers. Lined pattern forming upper part of volcano is subaerially erupted tholeiitic basalt, predominantly thin (2–10 m thick) pahoehoe and aa lava flows, but including ash layers. Lava from Mauna Loa predominates over that from Kilauea in north part of section. Dotted line indicates possible sole of future debris avalanche favored by steep regional slope.

Figure 8.39. A cross section showing the structure of the Hilina slump. Notice how this cross section resembles cross sections intended to show the shallowest section of subduction zones.

Landslides vs. Earthquakes or Cracks vs. Slip Pulses
Although the Hilina slump appears to be an example of an incipient long-run-out landslide, it has experienced other modes of failure. In particular, there have been several large earthquakes that seem to have occurred on the basal interface between the ocean floor and the volcanic deposits. The 1975 M 7.4 Kalapana

earthquake and the 2018 M 6.9 Leilani Estates earthquake both appear to have occurred as shallow-dipping thrust earthquakes along the interface at the base of the Hilina slump. Figure 8.40 shows the surface projection of the contours of slip that were used to simulate teleseismic P-waves.
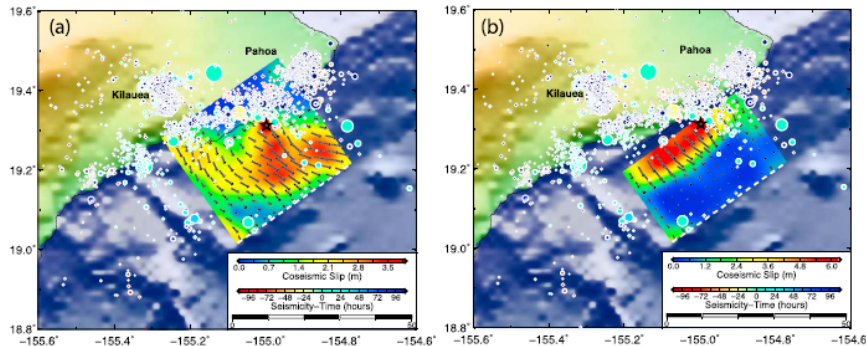

Figure 4. Finite-fault slip inversions using teleseismic broadband P and SH waves for the 4 May 2018 Hawaii Island earthquake for (a) $\phi = 235°$ and $\delta = 7.5°$, constrained to fit onshore GPS data and regional tsunami recordings (Bai et al., 2018), and (b) $\phi = 235°$ and $\delta = 3.0°$ constrained to fit the GPS data. The inversion in (a) gives $M_0 = 8.7 \times 10^{19}$ Nm (M_W 7.2), while that in (b) gives $M_0 = 7.5 \times 10^{19}$ Nm (M_W 7.2). U.S. Geological Survey catalog seismicity from 11 April to 14 May is shown by time-coded circles. Subfault slip direction and magnitude are indicated by arrows (direction of hanging wall movement). The black star indicates the 2018 event epicenter. The faults dip gently toward the northwest and are likely the basal décollement between the volcanic mass and the former Pacific seafloor. Inferred slip distribution and seismic moment depend on the dip angle estimate, so precise estimation is important.

Figure 8.40. Slip model of the 2018 M 6.9 Leilani Estates earthquake (Lay, T., Ye, L., Kanamori, H., & Satake, K.(2018). Constraining the dip of shallow, shallowly dipping thrust events using long-period Love wave radiation patterns: Applications to the 25 October 2010 Mentawai, Indonesia, and 4 May 2018 Hawaii Island earthquakes. Geophysical Research Letters, 45, 10,342–10,349. https://doi.org/10.1029/ 2018GL080042)

As is the case for almost all shallow crustal earthquakes, these moderate-sized earthquakes were probably propagating slip pulses. Furthermore, the dynamic friction during the slip pulses was probably very low (just like most shallow crustal earthquakes). The main difference between these earthquakes and a giant long-run-out landslide is that the landslides were certainly expanding cracks. Of course, a standard crack has a crack tip at at least one end (In many cases the crack extends to the surface). In the case of a landslide, both ends extend to the surface.

Presumably, there is a critical dimension for slip pulses. If the slip pulse becomes larger than this dimension, then it transitions into an expanding crack. At this point, I don't know what this critical size is, but I can make an educated guess. For example, the M 8.8 2010 Maule earthquake was well recorded by high-rate GPS stations located directly above the rupture. Figure 8.41 shows the displacements at several stations located above the rupture. Clearly this is an example of a slip pulse; the slipping had already stopped at stations in the south near the epicenter before rupture even began beneath stations in the north. Heaton, T., S. Minson, M Simons, 2011, S11A-2201 Characterization of the Slip Pulse for the 2010 M 8.8 Maule Earthquake, Fall AGU).

Fig. 8.41. GPS displacements recorded during the M 8.8 2010 Maule earthquake. This is clear evidence that the rupture consisted of a slip pulse; not all of the points on the rupture surface are slipping simultaneously.

While almost every reasonably well recorded earthquake seems to be best modeled with slip pulses, there are several examples that suggest that slip pulses grew so large that they turned into expanding cracks. I think that the best example of this behavior is from the 2011 M 9.2 Tohoku earthquake. Figure 8.42 shows the surface projection of contours of slip on the subduction interface (Simons, Minson, Sladen, Ortega, Science,). This slip model was constructed to simulate co-seismic GPS displacements combined with tsunami waveforms, which are primarily determined by the spatial pattern of coseismic uplift. Although the inferred slip is large (as much as 40 m in the hypocentral region), the inferred rupture history can be interpreted as a very large slip pulse. The Simons and others model appears to show that the rupture did not extend all of the way to the ocean floor. Unfortunately, the model resolution in the shallowest region (at the trench) is poor since the displacement data is only available from on-shore GPS stations. Furthermore, in the limiting case of zero depth, there are no radiated seismic waves for the case of a horizontal fault. That is, slip that occurs at the upper edge of a subduction earthquake is practically invisible for seismic studies.
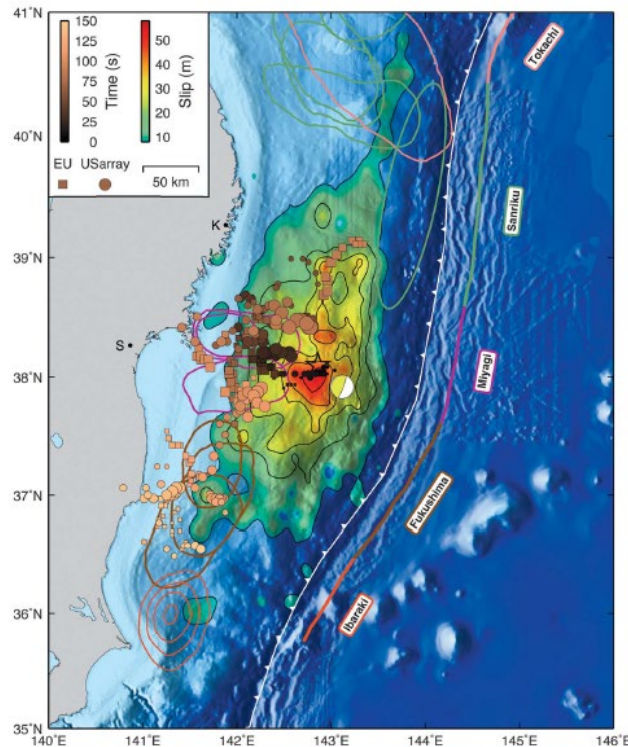
Figure 8.42

There was a lucky break for researchers studying the Tohoku earthquake. Ocean bottom observatories TJT1 and TJT2 had been installed just west of where the mega-thrust fault intersects the ocean floor (see Figure 8.43). Sonar positioning had been used to calculate the position of these stations on the ocean floor. Although the pressure records from the earthquake were of interest, the most amazing observation came from relocating the stations after the Tohoku earthquake. In particular, the change in position of TJT1 and TJT2 was determined to be 62 ± 20 m and 53 ± 20 m, respectively. In addition, the pressure gauges indicated that the stations had been uplifted by 5 m (Ito and others, 2011). It seems as though there was exceptionally large slip on the very shallowest part of the rupture. Ito and others showed that this large shallow slip was consistent with a large pulse in tsunami height that was recorded on two ocean bottom pressure sensors, TM1 and TM2 which are located closer to the Honshu coast (see Figure 8.43). Ito and others summarize their findings in Figure 8.44; they concluded that the slip in the 50-km wide shallow part of the rupture (average depth of 4 km from the ocean bottom) was about 80 m.

Remarkably, there was no evidence for this enormous slip to be found in the short-period accelerograms recorded on either Honshu or in the teleseismic body waves. Of course, this is what one would expect from crack-like rupture; the high-frequency radiation is only from the propagating crack tip. Furthermore, in the limit of a horizontal fault that becomes very shallow, there are no radiated waves of any sort (including surface waves). That is, very shallow thrusting on very low-angle planes is essentially invisible to seismometry. Unfortunately, this faulting geometry can cause static uplift of the ocean bottom, which can cause large damaging tsunamis. Shallow crack-like slumping at an ocean trench seems

8-78

to be a logical explanation for **tsunami earthquakes,** which are a class of nearly silent earthquakes (or slumps) that generate dangerous tsunamis (see McKenzie and Jackson, 2012, Tsunami earthquake generation by the release of gravitational potential energy, EPSL, 345-348, 1-8).
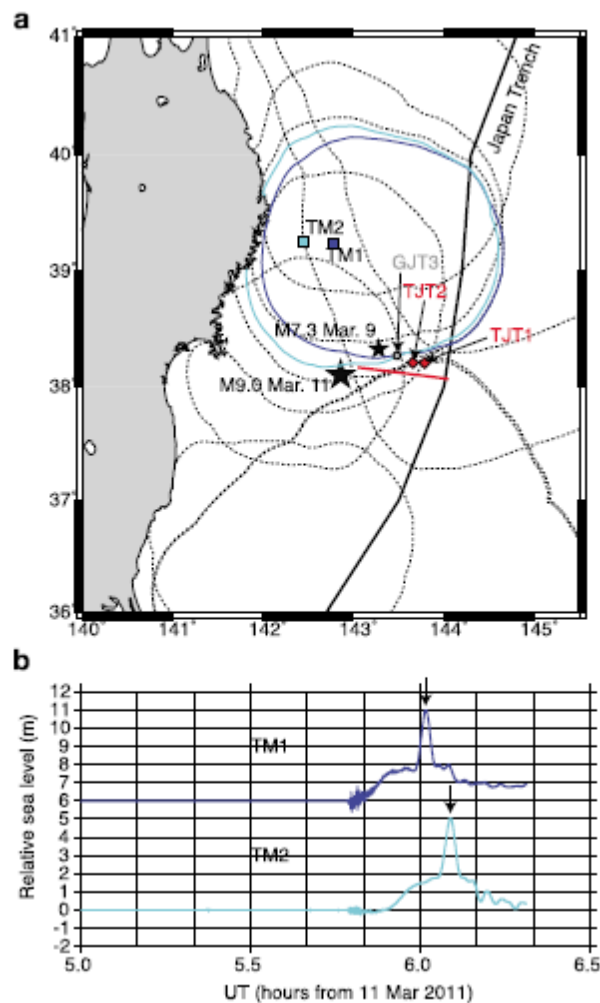

Figure 8-43. Locations of ocean-bottom observatories operating during the 2011 Tohoku earthquake (Ito, Y, T. Takeshi, Y. Osada, M. Kido, D. Inazu, Y. Hayashi, H. Tsushima, R. Hino, and H. Fujimoto, 2011, Frontal wedge deformation near the source region of the 2011 Tohoku-Oki earthquake, Geophys. Res. Lett., 38, L00G05, doi:10.1029/2011GL048355).
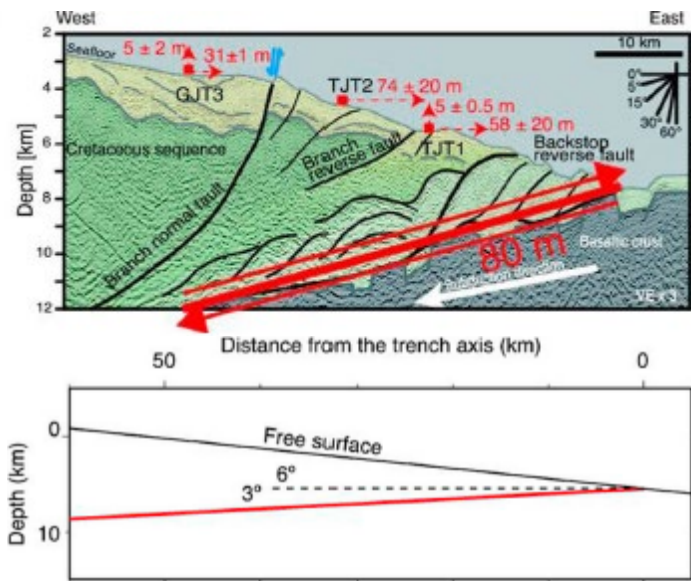
Figure 8-44. Schematic showing the region of inferred large slip near the Japan Trench for the 2011 Tohoku earthquake (Ito and others, 2011). The dip of the megathrust is so shallow that it appears that the shallowest part of the earthquake can be interpreted as a gravitational slump.

While it is uncommon to find direct evidence for crack-like rupture in earthquakes, crack-like rupture might also help to explain puzzling features of the enigmatic 1960 M 9.5 Chilean earthquake. Crack-like rupture produces very large slips that are spatially smooth; the ratio of short-period to long-period radiated energy is small for crack-like ruptures as compared to pulse-like ruptures of comparable potency.

Figure 8.45 shows the peak amplitude of broad-band (1 s to 90 s) teleseismic P-waves for the largest earthquakes of the 20th Century. M_W for these events is based on the potency (aka seismic moment) which was estimated from a variety of long-period measurements that include geodetic positioning, tsunami excitation, and excitation of free oscillations. Notice that, with the exception of the 1964 M 9 ¼ Alaskan earthquake, the largest earthquakes have long-period body waves that saturate with magnitude. This behavior is inconsistent with standard spectral scaling laws. It presents us with a mystery. Why are the potency rate functions of the largest earthquakes smoother than other earthquakes? Figure 8-46 shows potency-rate functions inferred for these largest events. These time functions were derived from teleseismic P-waves recorded by the long-period Benioff seismometer in Pasadena, which has a response that is flat to velocity between 1 s and 90 s (see chapter 3). In the case of the largest earthquakes (M > 9), it appears that there is a large portion of the potency rate function that is not visible in the 1 to 100 sec teleseismic P-waves.

One plausible explanation is that many of these largest events may have sections with crack-like ruptures, which are characterized by large, spatially-smooth slips.
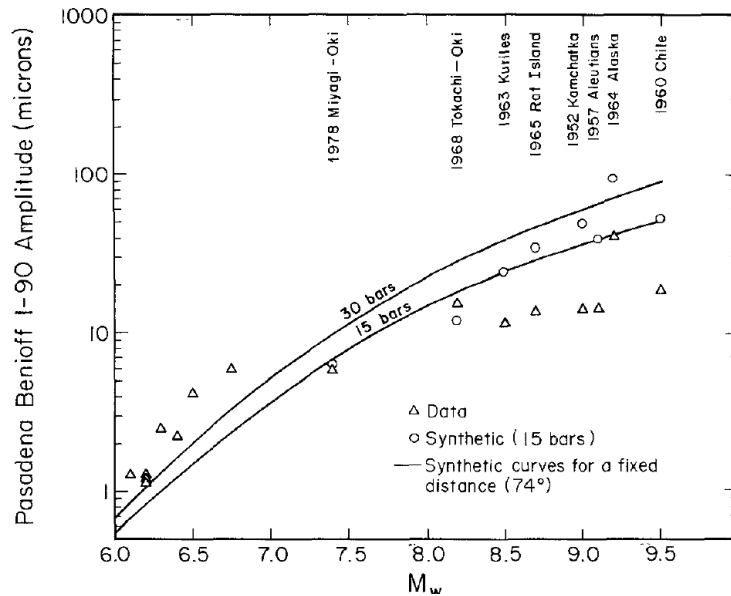
Figure 8-45. Maximum distance corrected P-wave amplitude on the Pasadena broad-band seismograph (aka, Benioff 1-90) for the largest earthquakes of the 20th Century. The triangles are observed data and the circles are predicted from standard f-2 spectral scaling. This scaling seems to match the observations of most earthquakes, except those with M > 9. Hartzell, S., and T. Heaton, 1988, Failure of self-similarity of large shallow subduction earthquakes, Bull. Seism. Soc. Am., 78, 478-488. http://resolver.caltech.edu/CaltechAUTHORS:20121121-115204715

Figure 8.46 shows Potency Rate functions that are 1) compatible with the observed 1-90 records, 2) have potencies that are compatible with estimates using free oscillation data, and 3) that have overall durations that are inferred from the source dimensions and typical rupture velocities (Hartzell, S, and T. Heaton, 1985, Teleseismic time functions for large shallow Subduction zone earthquakes, Bulletin of the Seismological Society of America , v. 75, pp. 965-1004. http://resolver.caltech.edu/CaltechAUTHORS:20130130-145105862). The apparent saturation of intermediate-period body waves is also recently reported by Kanamori and Ross (Reviving m_B, Geophys. J. Int. ,2018, **216,** 1798–1816 doi: 10.1093/gji/ggy510).

The bottom line seems to be that there is evidence that slip pulses can transition into growing cracks if the pulse becomes large enough (perhaps slips > 20 m are predominantly crack-like).
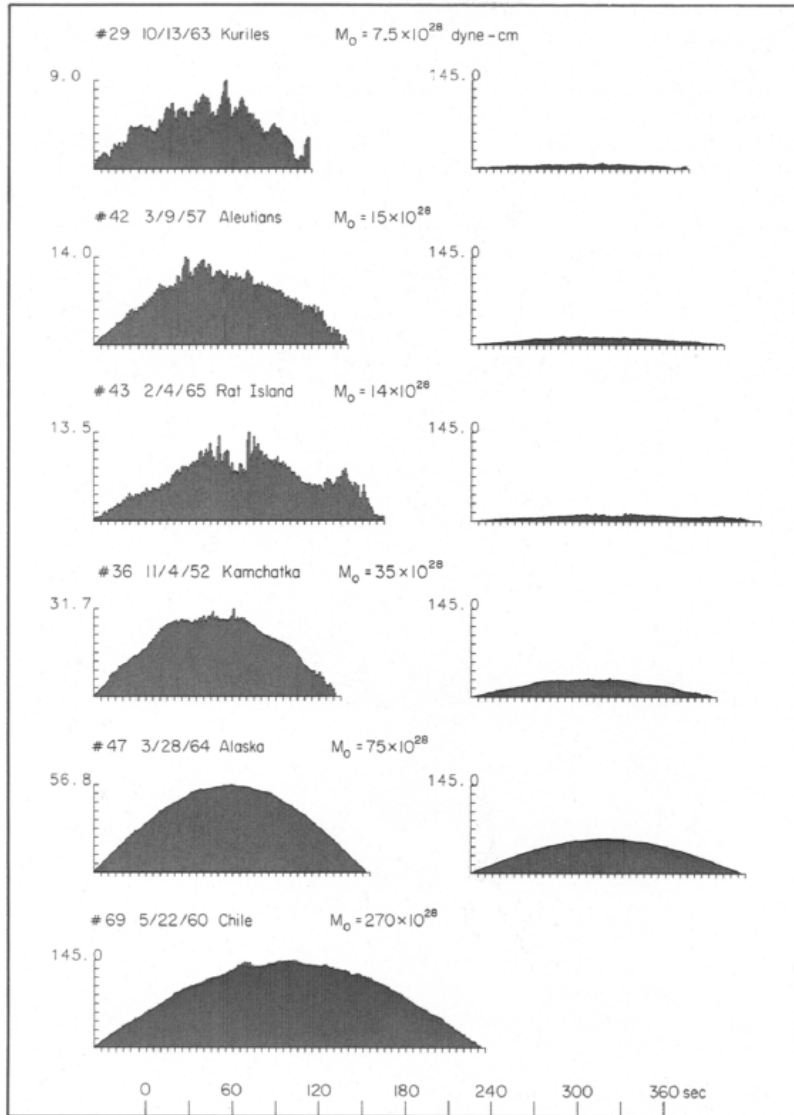
Figure 8. 46

Initiation of cracks

At the other end of the scale, earthquake nucleation is almost certainly a crack-like process. In order to have dynamic instability (an earthquake), it is necessary that the rate at which potential energy (usually strain energy) is removed from the system must exceed the rate of energy dissipation in inelastic processes (friction). For example, see equation 8.95, which estimates the minimum area that a crack-like rupture can experience dynamic instability. When Dieterich-Ruina friction is assumed for slowly increasing shear stress (e.g., plate motion), then inelastic strains tend to concentrate in an area that eventually becomes the hypocenter of an expanding crack. This temporal evolution is a consequence of the evolving state variable, and the nucleation process may accelerate over days in a small patch. The important point is that ruptures cannot initiate as slip pulses; the rupture area of the slip pulse would be smaller than the critical length for instability.

Transition from cracks to pulses

Assuming that all earthquakes initiate as growing cracks, and that most significant shallow crustal earthquakes appear to be propagating pulses, leads to the conclusion that there must be some length where there's a transition from crack-like rupture to pulse-like rupture. Currently, finite fault modeling of seismograms is the methodology that allows us to image slip pulses. This detailed 3-d imaging is more difficult when the distance to the observation is large compared with the dimension of the rupture. That is, it has not been feasible to image the ruptures of earthquakes smaller than M 5 ½ with enough precision to resolve slip pulses.

Another way to investigate the transition of cracks to pulses is shown schematically in Figure 8.47 (from Meier, Heaton, and Clinton, 2016). This schematic shows several circular arcs on a rupture plane. The arcs are meant to show crack (or healing) fronts at different times. The shaded areas show the slipping region for several different times. For areas near the point marked as nucleation, all of the area is slipping in the moments after the rupture front has passed; it's an expanding crack. At later times, the rupture front has propagated further from the hypocenter and the slip is now confined to a pulse.

The schematic shows expected slip profiles at different times. When the earthquake is crack-like, we expect to see spatially smooth slip profiles, whereas when it's a pulse, we expect to see spatially heterogeneous slip; I'll come back to that to explain why pulses are inherently unsteady. Panel d is a schematic of the potency rate function that one would expect from such a conceptual model. Remember that the far-field displacement should look like the potency rate function.
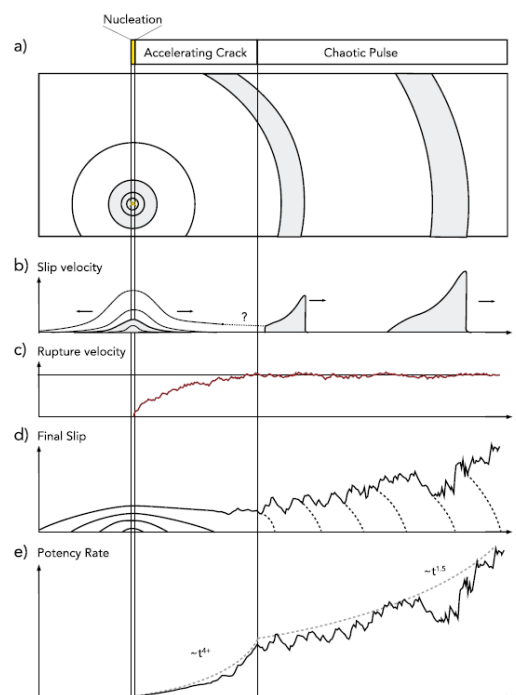

Fig. 8-47 Cartoon showing concept for transition between crack-like and pulse-like rupture (see text).

Meier, Heaton, and Clinton investigated a large set of near-source displacement records (doubly integrated from accelerograms). The records were all at hypocentral distances of less than 25 km and they were all processed in the same way. The records are sorted into magnitude bins and then the log of the median peak displacement for each magnitude bin is plotted as a function of log time (see Fig. 8-48). The use of a log-log plot allow us to look for power-law behavior. Time is defined with respect to the first P-wave. Pay close attention to the logarithmic time scale; the first decade is only one tenth of a second and the center of the plot is just 1 second. The S-wave pick times are displayed at the bottom of the plot as single points for each record. Notice that the S-waves do not begin to arrive until several seconds after the P-arrival. Also notice that for the first 0.2 s, the median motions are approximately the same, independent of the magnitude. Furthermore, the pgd's grow quickly and increase as approximately t^3. We interpret this initial rapid growth as crack-like rupture in the hypocentral region. There also appears to be a change in the pgd growth rate at about ¼ s. That is, pgd seems to grow as t^3 until it transitions to a growth rate of about $t^{1.3}$ for times greater than ¼ s. Of course, earthquakes smaller than M 4.5 seem to be all over by ¼ s, so the transition to $t^{1.3}$ growth is only observed for earthquakes larger than M 4.5.

Fig. 8.48

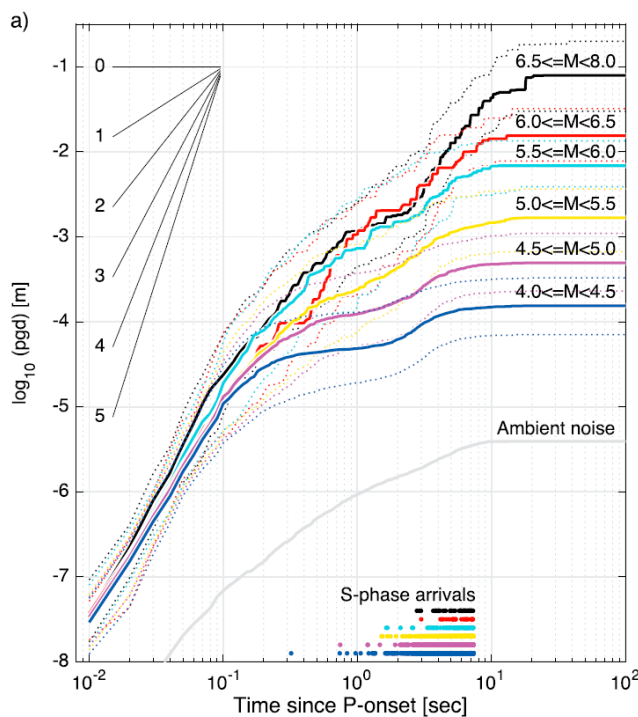I interpret this break in growth at 0.25 s (corresponds to M 4.5) as the transition from cracks to pulses. I admit that this evidence is suggestive and not compelling. If this is the transition from cracks to pulses, then we can estimate the rupture length at which the transition occurs. For crack-like rupture, the total event duration is about twice the time required for the rupture to spread, or

$T_{event} \approx 2\sqrt{S}\big/V_R$. $L_{crack-pulse} \approx \frac{1}{2}V_R(0.25s)$. If we use a typical rupture velocity of 2.5 km/s, then the transition would occur at a rupture length on the order of 250 m. Unfortunately, the assumption of a length-scale independent rupture velocity is not at all obvious. That is, it seems plausible that the rupture velocity increases as the crack grows. That possibility of increasing rupture velocity seems difficult to assess and I will ignore it for the present. Assuming that the dimension of the rupture at which this break in scaling is 250 m, then the rupture area would be on the order of 6×10^4 m^2 (assuming that the rupture approximately square). We can use 8.22 to estimate that that corresponds to a magnitude 3.0 earthquake. This estimate is similar to the M 4.5 at which the scaling changes in Fig. 8.48. Using 8.30, we can estimate the total change in potential energy as

$$\Delta E_W \approx SD\left(\sigma_0 - \frac{1}{2}\Delta\sigma\right)$$

$$\approx 3 \times 10^{-5} S^{3/2}\left(\sigma_0 - \frac{1}{2}\Delta\sigma\right) \qquad 8.106$$

Or,

$$\frac{\Delta E_W}{S} \approx 3 \times 10^{-5}\sqrt{S}\left(\sigma_0 - \frac{1}{2}\Delta\sigma\right) J\Big/m^2 \qquad 8.107$$

If the rupture area is 6×10^4 m^2, then $\dfrac{\Delta E_W}{S} \approx 2.4 \times 10^{-2}\left(\sigma_0 - \frac{1}{2}\Delta\sigma\right) J\Big/m^2$. If

the prestress σ_0 is a typical Byerless stress of 200MPa, and if the stress drop is the average of 3 MPa. Then the change in potential energy of a M 3.0 is about 5 MJ/m^2. This enough energy to melt several mm's of fault material. This potential energy change can be compared with the radiated energy that is estimated using 8.57. The following is from an unpublished note from Hiroo Kanamori. I lightly modified his equations which were written in the familiar form of seismic moment and assuming a circular rupture. I will assume the standard f^{-2} model

$$E_R(f) = 8\pi P^2 f_c^4\left(\frac{\mu^2}{15\rho\alpha^5} + \frac{\mu^2}{10\rho\beta^5}\right)\int_0^f \frac{f^2}{\left(f^2 + f_c^2\right)^2}\,df$$

$$= 8\pi P^2 f_c^4\left(\frac{\mu^2}{15\rho\alpha^5} + \frac{\rho}{10\beta}\right)\int_0^f \frac{f^2}{\left(f^2 + f_c^2\right)^2}\,df \qquad 8.108$$

$$= 8\pi P^2 f_c^4 \rho\left[\left(\frac{\beta}{\alpha}\right)^4 \frac{1}{15\alpha} + \frac{1}{10\beta}\right]\int_0^f \frac{f^2}{\left(f^2 + f_c^2\right)^2}\,df$$

The S-wave energy is large compared to the P-wave, so neglect the P-wave.

$$E_R(f) \approx 0.8\pi P^2 f_c^4\, \rho\Big/\beta \int_0^f \frac{f^2}{\left(f^2 + f_c^2\right)^2}\,df$$

$$= 0.8\pi P^2 f_c^3\, \rho\Big/\beta\left[-\frac{ff_c}{2\left(f^2 + f_c^2\right)} + \frac{1}{2}\arctan\left(\frac{f}{f_c}\right)\right] \qquad 8.109$$

If $f_1 \gg f_c$, then

$$E_R\left(f_1 < f < \infty\right) \approx \frac{2\pi\rho P^2 f_c^3}{5\beta} \frac{f_c}{f_1}$$

8.110

Using the standard Brune model,

$$f_c \approx 0.42\beta\left(\frac{\Delta\sigma}{\mu P}\right)^{1/3}$$

8.111

So

$$E_R\left(f_1 < f < \infty\right) \approx \frac{2\pi\rho P^2 C^3 \beta^3 \left(\frac{\Delta\sigma}{\mu P}\right)}{5\beta} \frac{C\beta\left(\frac{\Delta\sigma}{\mu P}\right)^{1/3}}{f_1}$$

$$= \frac{2\pi\rho P^2 C^4 \beta^3}{5f_1}\left(\frac{\Delta\sigma}{\mu}\right)^{4/3}$$

8.112

$$\approx 0.04\rho P^2 \beta^3 \left(\frac{\Delta\sigma}{\mu}\right)^{4/3} \frac{1}{f_1}$$

To compare the radiated energy per unit rupture area with the potential energy per unit rupture area, we need to divide 8.112 by the rupture area

$$S = \left(\frac{7\mu P}{16\Delta\sigma}\right)^{2/3}$$

8.113

And we obtain

$$\frac{E_R\left(f_1 < f < \infty\right)}{S} \approx 2\times10^{-2}\frac{\Delta\sigma^2}{\rho\beta f_1}$$

8.114

If we assume that $\Delta\sigma = 3\text{MPa}$, $\rho = 3\times10^3 \text{ kg}/\text{m}^3$, and $\beta = 3.5 \text{ km}/\text{s}$, then $\frac{E_R\left(f_1 < f < \infty\right)}{S} \approx 20\frac{1}{f_1}$ kJ/m^2.   However, the 3 MPa stress drop is a number that is appropriate to simulate static stress drop. Our current derivation is designed to simulate near-source high-frequency motion. In this case, it seems more appropriate to use effective stress instead of static stress drop. Hanks and McGuire used near-source acceleration records to conclude that effective stress is about 10 MPa. Inserting 10 MPa into 8.114 gives an estimated radiated energy of 2MJ/m^2. Clearly, there is ambiguity in obtaining the radiated high-frequency energy estimate. However, I would say that it's in the range of 1 MJ/m^2.

In order to form a slip pulse, there must be enough potential energy change to sustain the high-frequency radiated energy in a slip pulse. If we say that the high-frequency energy (f > 2 Hz) radiated by a slip is described by a white noise

acceleration spectrum (this is basically Brune's model), then we can use 8.114 to provide a very rough estimate of the radiated high frequency energy. Since the rupture area of a slip pulse is small compared to a crack, the estimate may be larger than that just provided (as large as 200 kJ/m^2??). Our earlier estimate of potential energy change of 5 MJ/m^2 for a M 3 indicates that there is abundant available energy to support the radiation of high frequency waves (greater than 2 Hz) that are observed in the near-source region of slip pulses.

The Sound of Slip Pulses

I have used the phrase, "white noise," without providing much definition or context. The *"colors of noise"* is a concept that originates in audio engineering. It has been observed that many forms of noise have random phase with a power spectrum that is approximately described by a power law $f^{2\alpha}$. Recall that the power spectrum is an energy spectrum. That is, if the power spectrum is $f^{2\alpha}$, then the amplitude spectrum is f^{α}. In the case of sound, the amplitude of sound is proportional to acoustic pressure, which is proportional to the particle velocity.

In audio engineering, the term, white noise, refers to a sound that has equal energy in all frequency bands ($f^{2\alpha=0}$, just like white light). In this case, the amplitude spectrum is also flat and white. A good example of the sound of white noise is the sound that an old-fashioned AM radio makes when it's tuned between stations (sometimes called shot noise). If the amplitude spectrum of the particle velocity is constant, then the amplitude spectrum of the particle displacement is f^{-1}, and the spectrum of the particle acceleration is f^{1}. Now the observation of the "sound of a slip pulse" is that it has constant particle acceleration amplitude, or the velocity amplitude spectrum is f^{-1}. This in turn gives a power spectrum that is f^{-2}. Noise of this type is typically called red noise. This form of noise is quite common in physical systems. For instance, the time integral of randomly timed force impacts (e.g, Brownian motion) produces f^{-1} red noise. Red noise sounds very much like the roar of a waterfall.

It seems that slip pules are accompanied by white-noise acceleration with a total radiated energy on the order of 1 MJ/m^2. It's as if slip pulses have a distinctive sound (the sound of a waterfall). Furthermore, this noise seems to be consistent with the notion that the slip pulse is characterized by Brownian-motion-like impacts from a fluidized granular material.

Modeling Slip Pulses

In my 1990 paper, I presented observational evidence that earthquake rupture is pulse-like and I then presented a conceptual model of dynamic friction that I claimed would produce slip pulses. In particular, I hypothesized that dynamic friction was inversely proportional to slip speed, which is a particular realization of the friction law in equation 8.99. At the time that I wrote the slip pulse paper,

I searched for an analytic solution for a steady slip pulse. The only solution that I found is a solution by Freund (1970). Fig. 8.50 shows Freund's solution for a mode III slip pulse of length l_p that is propagating at constant velocity V_R. The shear stress at infinity is uniformly equal to τ_0 and the shear stress on the slipping patch is τ_p. The dynamic stress drop of the pulse $\Delta\sigma_p$ is given by

$$\Delta\sigma_p \equiv \tau_0 - \tau_p = \frac{\mu}{\pi}\frac{D}{l_p}\sqrt{1-\left(\frac{V_R}{\beta}\right)^2} \qquad 8.115$$

Which implies that

$$D = \Delta\sigma_p l_p \frac{\pi}{\mu}\frac{\beta}{\sqrt{\beta^2 - V_R^2}} \qquad 8.116$$

Or alternatively,

$$l_p = \frac{D}{\Delta\sigma_p}\frac{\mu}{\pi}\frac{\sqrt{\beta^2 - V_R^2}}{\beta} \qquad 8.117$$
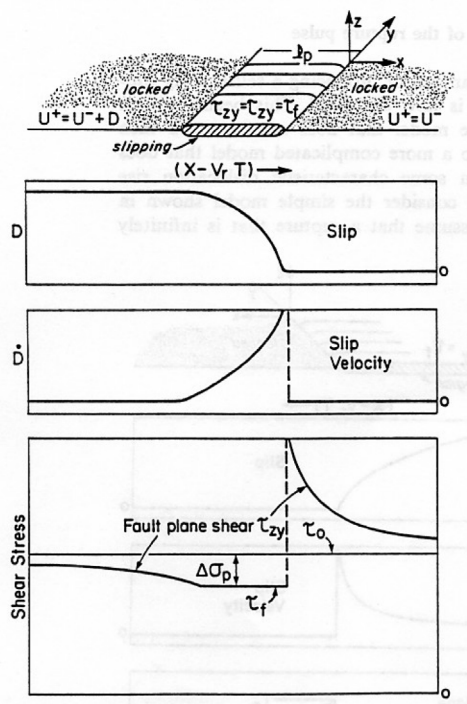
Figure 8-49. Idealized model of Freund (1970) in which a pulse of slip propagates steadily at velocity V_R. A uniform shear stress τ_0 is applied at infinity and the shear stress on the slipping portion of the fault is assumed to be to τ_f .

This solution is obtained by the appropriate transformation of the static solution of a mode III crack of length l_p . This solution has numerous quirks that limit it's usefulness. In particular, if the rupture velocity approaches the shear

velocity, then, for a given slip, the length of the slip pulse tends to zero. Or alternatively, given the length of the slip pulse and the slip, then the dynamic stress drop of the slip pulse tends to zero as the rupture velocity approaches the shear wave speed. These features are the result of singular stress at the crack tip that results from instantaneous stress drop (no fracture energy). Presumably, the stress at the leading edge of the slip pulse is approximately the shear stress that is required to overcome static friction.

As was mentioned earlier, there is no inelastic work at the crack front (fracture energy) and the crack front has singular stresses that cannot be handled numerically. The second issue is that this friction law has no stable steady-state solutions. In particular, a steady-state pulse should propagate at a constant velocity and with a slip pulse of constant slip.

Another strange aspect of this solution is that the static stress drop is zero. That is, the pulse propagates indefinitely, so the total rupture length is infinite while the slip is a constant. Finally, since this solution is obtained from a transformed static solution, there are no radiated waves generated by this slip pulse. This turns out to be an important **feature of a slip pulse that propagates steadily. It is invisible to seismology. A slip pulse only radiates far-field waves when 1) either it changes its amplitude as it propagates, or 2) it changes its rupture velocity.** You can obtain further insight in this by reading the analysis associated with Figures 7.12 and 7.13. Interestingly, this invisibility of steady-state slip pulses means that we **cannot resolve the upper limit of slip velocity from seismic data.** That is, you can only observe the details of a slip pulse if you are so close to the pulse that you can observe the near-field terms (basically the static part of the solution). Since seismic measurements are only available from the Earth's surface, it is practically impossible to observe the true nature of slip pulses that propagate at depth greater than a couple of km. . However, it is possible to put a lower bound on the average slip velocity. In particular, the duration of pulse-like ground velocities in the forward directivity direction provides a lower bound on slip velocity. Since directivity pulses are only observed for mode II ruptures (they are on the fault normal component of stations located near a fault with the provision that the rupture is is propagating parallel to the slip direction (i.e., mode II). Modeling of records of this type indicates that the length of slip pulses is typically less than 10% of total rupture length.

It appears that the average slip velocity in slip pulses is greater than 2 m/s. Since the elastic strain in shear waves is on the order of the particle velocity (1/2 the slip velocity) divided by the wave speed (see Chapter 3), we can conclude that a particle velocity of 1 m/s in a material with a shear wave speed of 3 km/s implies a shear strain of about 3×10^{-4}. Therefore, a slip velocity of 20 m/s implies a shear strain of about 3×10^{-3}, which is about the maximum shear strain that rocks can sustain at the confining pressure of earthquakes. While 20 m/s slip velocities (at 10 km depth) is an order of magnitude faster than is typically assumed for earthquakes**, it is not possible to exclude the possibility of slip velocities as high as 20 m/s**. Again, high slip velocity in a slip pulse that is steadily propogating at depth does not necessarily produce large high-frequency radiation.

It is important to recognize that the solution given by 8.115 is not a "self-healing" pulse. It is only a slip pulse because I constrained the fault rupture to be pulse-like. In my 1990 slip-pulse paper, I hypothesized that friction that scales with inverse slip velocity should produce slip pulses. As it turns out, it has not been possible to simulate the behavior of two half spaces that are governed by a friction law that varies inversely with slip speed.

While a solution like that shown in Fig. 8.24 can match boundary conditions, it is not a stable solution. In order to achieve a stable solution, then it must be that slight perturbations from homogeneity in the prestress should have little effect on the solution. Assume that you have found a slip pulse that will propagate steadily at a particular homogeneous prestress. Now assume that some slight perturbation occurs as the pulse runs (e.g. a tiny patch with higher prestress). When the pulse hits the perturbation, then the pulse temporarily begins to grow larger, which means that it has slightly larger slip speed. However, larger slip speed means that the friction drops, which causes the pulse to grow even larger. That is, if you put a tiny positive perturbation to a steadily running slip pulse, then it grows spontaneously without bound. Similarly, if you perturb the solution such that it temporarily becomes smaller, then it quickly shrinks to zero.

Rice and Lapusta recognized this difficulty and they concluded that a friction that is purely rate weakening would produce slip pulses with unstable solutions (Rice, J. R., Spatio-temporal complexity of slip on a fault, J. Geophys. Res., 98(B6), 9,885–9,907, 1993). They concluded that pure rate weakening friction is a pathological mechanics problem. They said that this problem is ill posed. By that they meant that there was insufficient information to construct a solution to the problem. In particular, there is no length scale to rate weakening friction; while a stable slip pulse has a length scale associated with it (see Chapter 5 of Ahmed Elbanna's PhD dissertation for further discussion).

Aagaard and Heaton (2008, Constraining fault constitutive behavior with slip and stress heterogeneity, J. Geophys. Res., 113, B04301, doi:10.1029/2006JB004793) studied this problem using 3-d finite element models of sliding faults with a friction law that had both slip weakening (fracture energy) and rate weakening. They were unable to simulate rupture using pure rate weakening (i.e. zero fracture energy) since the solutions became singular at the crack tips; it was not possible to attain numerical stability. Based on this knowledge, they hypothesized that, if there is a solution to pure rate-weakening friction, then the solution would have to be described by a fractal, since fractals and homogeneous solutions are the only classes of solution that have no inherent scale.

Aagaard and Heaton were searching for ways to produce earthquakes that had self-sustaining heterogeneity. That is, they observed that most finite fault source inversions resulted in slip models that 1) were characterized by slip pulses, and 2) the slip models were spatially complex. Aagaard and Heaton attempted to find ways to produce self-sustaining spatial heterogeneity that would persist through many millennia of earthquakes. Although they were able to find spatial distributions of frictional properties that would produce spatially heterogeneous slip in response to plate motions. They found that the residual stress that was left after the heterogeneous event was then spatially correlated with the heterogenous

friction. The result was that the next event in the sequence was spatially homogeneous. Aagaard and Heaton wanted to show that strong rate weakening could produce earthquakes that were 1) self-sustaining heterogeneous, 2) skip pulses, and 3) that required very little inelastic dissipation energy (friction and fracture energy). Unfortunately, achieving numerical convergence with low fracture energy and with strong rate weakening was not possible with even the largest super computers. When they recognized that the number of computations required for this calculation grows as $\frac{1}{\Delta x^4}$, they realized that finite element simulations could not solve this problem. Worse, the simulation of strong rate weakening and low dissipation requires that the model should evolve into a prestress that is spatially complex. In particular the prestress must be fractal (I'll get back to that later). This makes the numerical simulation problem even more overwhelming.

Aagaard and Heaton found that they could produce more spontaneous slip heterogeneity by increasing the strength of the rate-weakening in their friction law. A general observation that comes from numerical finite element modeling of rupture dynamics is, if you assume relatively homogeneous prestress, then the change in potential energy per unit rupture area grows as the rupture length grows. The only parameter that is available to absorb all the potential energy is the fracture energy; the fracture energy must exceed the radiated energy if you want to control the dynamics of an earthquake. I have seen numerous dynamic models with enormous fracture energies. My opinion is that these reported fracture energies are simply introduced as a way to avoid numerical instability in the calculation of stress at the rupture fronts. I will shortly argue that prestress is extremely heterogeneous and that this prestress heterogeneity serves to limit the growth of dynamic ruptures.

Following the realization that there are fundamental difficulties with 3-d finite-element modeling dynamic ruptures assuming strong rate-weakening friction, Ahmed Elbanna and I took a different approach. We decided to study the impact of strong rate weakening friction on a 1-d spring-block-slider model (sbm). The sbm (shown in Fig. 8.50) consists of a number $(i = 1,...,m)$ of identical point masses that are connected with identical leaf springs (stiffness k_l) to rigid blocks (above and below) that move with a steady differential velocity v. Additionally, the blocks are connected to each other by identical coil springs (stiffness k_c). Finally, there is uniform rate-weakening friction between the masses and the lower rigid block ($F_{fi} = \dfrac{1}{1 + b\dot{u}_i}$). Dynamic friction is specified to occur at any block that exceeds a specified maximum force. While you might think of this maximum stress in a block is the "strength" of this system, it seems that changing the value of this maximum has very little effect on the overall behavior of this system.
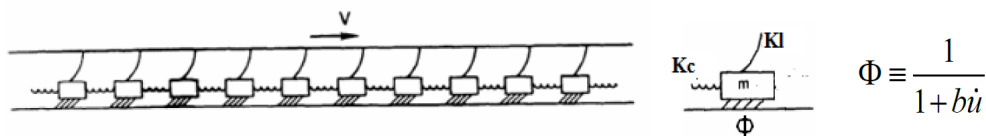
$$\Phi \equiv \frac{1}{1+b\dot{u}}$$

Figure 8.50

The sbm has several tunable parameters that control the system behavior. In particular, if the coil springs are stiff, then there is strong coupling between the masses and when failure occurs, it tends to involve all of the blocks. In fact, events tend to be crack-like in the sense that all bocks are moving during an event. As I have already mentioned this friction law provides a positive feedback in the dynamics of this problem. That is, the faster the sliding, the lower the friction which produces even faster sliding. Positive feedback dynamics is often a characteristic of chaotic systems. For example, the stock market is a highly complex system. When stock prices are rising, people want to buy so they can get in on the profits. Conversely, when stock prices are falling, people want to sell before they lose their investment. Wild fires are another example of a positive feedback system. The larger a fire, the faster it spreads. Epidemics are another example of a positive feedback system. All of these systems have complex behavior that is difficult to predict. In fact, often the behavior is so complex that we resort to random variables to describe them. However, it is important to recognize that complexity does not necessarily imply random. In fact, surprising structure can spontaneously appear as the result of chaotic dynamics.

Elbanna and I chose to investigate the sbm model since it is simple enough to explore with modern computers, while it still has characteristics that are of interest for earthquake dynamics. In particular, it is possible to explore the consequences of strong rate weakening friction using an sbm model. One reason that sbm's are amenable to computer simulations is that they are nearest-neighbor models. That is, motion of a single block only affects the adjacent blocks. This is in contrast to an elastic continuum; slip on a fault segment causes stress changes throughout the entire medium. The fact that the motions at a finite-element node are affected by the motions of all other nodes is responsible for the enormous size of 3-d continuum calculations. By their very nature, slip pulses are short-range phenomena, and it's critical to have a numerical scheme that concentrates on short-range phenomena if there is any hope of studying them.

There are two distinct stiffnesses in the sbm, the coil springs and the leaf springs. If the friction is dropped to zero, then the coupled blocks vibrate through traveling waves. These waves are dispersive because of the two stiffnesses. If the system is frictionless and composed of only coil springs, then the system is nondispersive and the sound wave velocity is, $c_s = \sqrt{k_c / m}$. If leaf springs are added to the frictionless system, and if it is excited by harmonic forcing of frequency ω, then the wave speed is dispersive and is given by

8-92

$$c_p = c_s \sqrt{1 - \frac{k_l}{(\omega c_s)^2}}$$ 8.118

Fig. 8.51 shows the spatio-temporal characteristics of a typical slip pule in an sbm with velocity weakening friction. Notice that the pulse is very localized as it runs (pulse width is 5% of the rupture length). Furthermore, the pulse seems to be nondispersive and running at the sound velocity (since the springs are linear, there is no way for disturbances to jump to supersonic velocities). The apparent constant rupture velocity is typical of most events in this system and it is a key feature that allowed us to develop the slip-pulse energy equation that allows us to simulate this system with a single equation (more later on the pulse-energy equation).

This vertical distance repressnts the duration of the rupture at block 40. It is simply the difference in the arrival times of the healing tail and the rupture front at that time
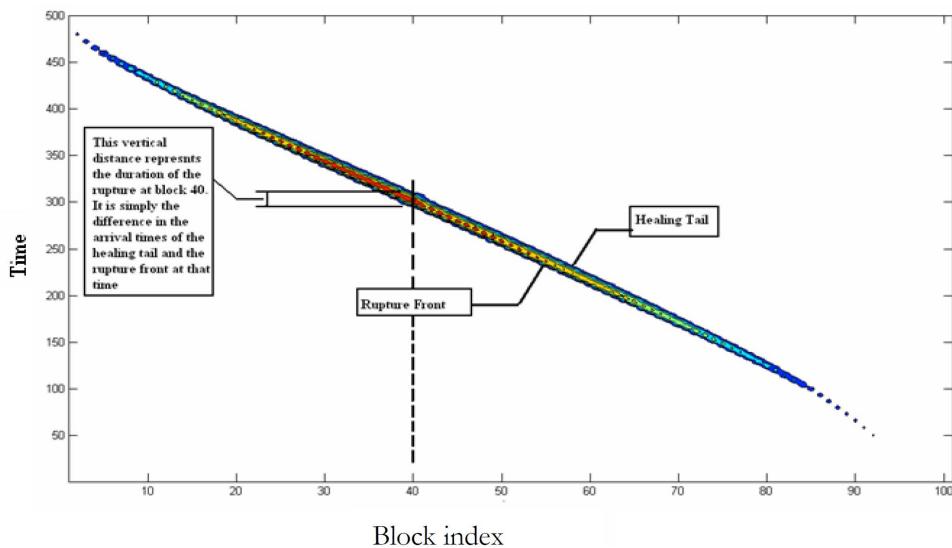
Healing Tail

Rupture Front

Block index

Figure 8.51

Even though all of the springs, masses, and friction laws are identical for every block in our model, the individual events are quite complex. Typically, we initiate the system by assuming a higher pre-stress in some group of blocks. Although individual events can become quite large, they do not grow without bound since the energy in the leaf springs that drives the system has an upper bound. After a large number of events have occurred, the system attains an unusual kind of steady state. That is, the statistical properties of the events and the prestress no longer change as more events occur. That is, the systems evolve towards a **strange attractor**. In dynamics, an attractor is a state that a system tends to evolve to independent of initial conditions. An attractor is called strange if is described by a fractal. An example of the nature of the prestress that has evolved after tens of thousands of events is shown in fig. 8.52.
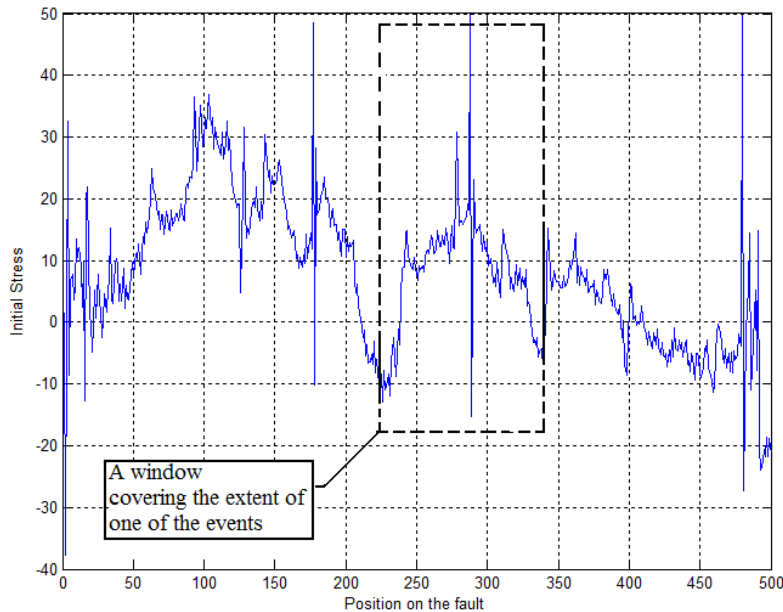
Figure 8.52

The model is unitless and the maximum shear before dynamic failure is 50 units. In this example, the next event starts as an expanding crack at the narrow spike at block 290 and it then quickly becomes two slip pulses that propagate bilaterally until they arrest at block 225 on the left and block 345 on the right. Notice that the pulses propagate at the relatively low prestress of 10 units and that they arrest after propagating across regions of **negative shear stress**. Many colleagues have been surprised that a frictional system that is being forced with right-lateral shear can evolve patches of left-lateral shear stress. In the case of slip pulses in an sbm, the answer is quite simple. As a slip pulse propagates, it can become accumulate large kinetic energy (i.e., large slips). Because of the strong feedback in the friction law, slip pulses can also arrest very abruptly. When this happens, there can be a patch of very large slip that is frozen in when the event is over. The slip may be locally large enough that it causes shear stress in the opposite direction of the motion of the rigid plates driving the system.

Even though there are patches of negative shear in this self-organized prestress, there are never any events that have slip in the opposite direction of forcing. That is because events only nucleate when the stress exceeds a threshold and this only happens when the prestress has the same polarity as the driving stress. This same principle applies to the Earth. If there are patches of negative prestress (e.g., left-lateral stress on the San Andreas fault), then they are mostly invisible to seismology since such patches will not nucleate future events. For more discussion on this subject, see Smith and Heaton.

Figures 8.53 shows an example of the types of events that occur after the system has evolved into a strange attractor. We call this example Case A. In each of these plots, a sequence of thousands of consecutive events are plotted. Each event (event index is the horizontal axis) shows the blocks that moved during that event. Fig. 8.53 shows events number 62,000 through number 66,000. Most of these 4,000 events involve only short ruptures, although there are many long

8-94

ruptures as well. If you stare at this sequence, it would be easy to convince yourself that there are likely to be some blocks with different properties that control the locations of events. Of course, this is not the case; everything is uniform in the model.

There are also persistent local knots that experience a large number of mostly short ruptures. These knots sometimes persist even after a long rupture extends through the knot. Inevitably, though, these knots are transitory, and some future large event erases them and causes new knots to appear. It's important to remember that sbm's only have nearest neighbor interactions. This means that once a knot forms, conditions at the knot are totally independent of what is happening at more distant nodes in the system.
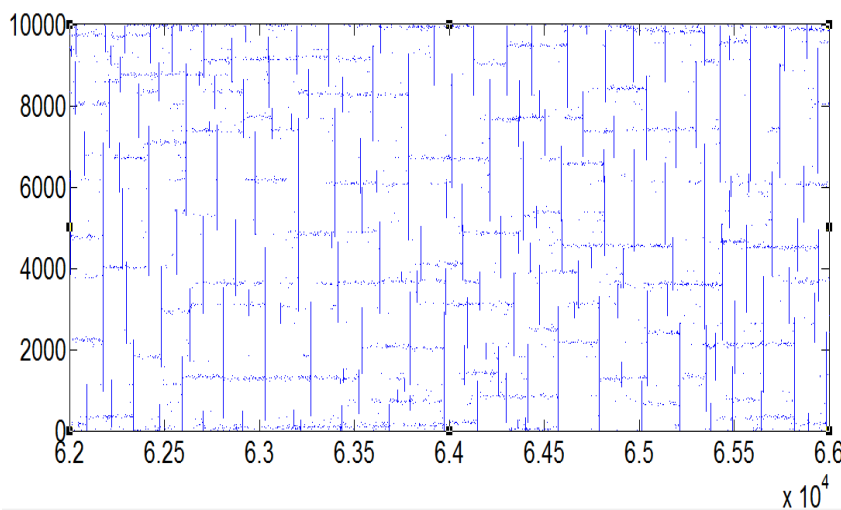

Figure 8.53

Although I claim that this system is chaotic, there is actually quite a bit of structure in the spatio-temporal sequence of the larger events. Complex does not mean that it's random. In nonlinear dynamics, a system is called chaotic if its long-term solution is sensitive to perturbations in initial conditions. If a small change in the prestress at some time causes the future events to diverge from the unperturbed system, then the system is chaotic, which is the case for this system.
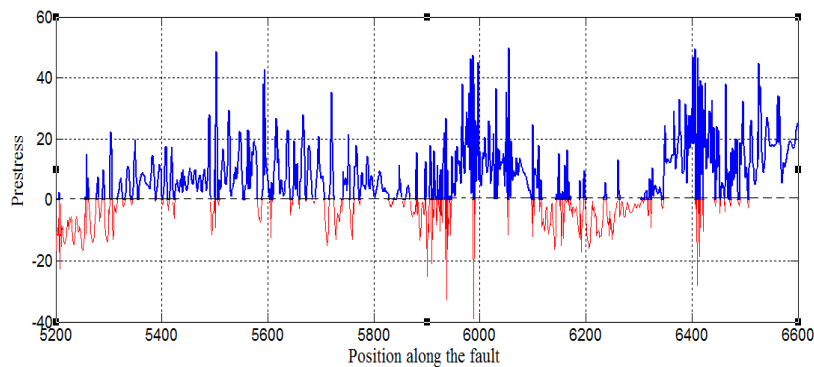

Figure 8.54

Fig. 8.54 shows a snapshot of the shear stress that is associated with Figure 8.53 (case A). Only 1,400 of the 10,000 total blocks are shown to better display the

spatial complexity. Areas where the prestress is negative (opposite to the driving stress) are shown in red. Even though the next event on this system is entirely controlled by this prestress distribution, it is virtually impossible to guess what the next event will look like. You've got to run the computer to calculate the complex evolution of the next slip pulse as it propagates through the system. Later, I will show an equation that does in fact do this calculation. An important feature of the system is that two events with virtually identical prestresses can be very different if the events nucleate at different places.

Figure 8.55 shows the relationship between average slip in an event and the rupture length of the event for Case A. Although there is large scatter, there is a clear trend that shows that events with larger rupture lengths have larger average slips. In fact, the ratio of the average slip divided by the rupture length is approximately independent of the overall size of the event (as measured by the rupture length). Figure 8.48b shows the same events, but in this case, the vertical axis is the average change in stress (stress drop) and the horizontal axis is rupture length. This implies that the average stress drop in this system is independent of the event size. This is very similar to the observations of natural earthquakes. In the case of chaotic slip pulses, the **scale invariance of stress drop is a consequence of self-organization of the prestress**; it's not caused by details of the friction law.

Although the stress drop is scale invariant in this model, there is still a scatter of about a factor of 100 between individual events. Similar scatter is seen in natural earthquakes. I have the impression that many researchers believe that the scatter in observed stress drops is primarily caused by erroneous measurements. Personally, I believe that most of this scatter is real and that it comes from the natural variation arising from rupture complexity of slip pulses. In fact, the average stress drop of earthquakes is almost certainly a measure of the heterogeneity of the spatial distribution of slip. That is, if the amplitude of a slip pulse changes slowly as it propagates, then the probability that the pulse will drop to zero is small compared to a similar sized pulse whose amplitude changes rapidly as it propagates. This topic is discussed in a paper by Liu-Zeng, Heaton, and DiCaprio, 2005, The effect of slip variability on earthquake slip-length scaling, Geophys. J. Intl., 162 (3), 841-849. http://resolver.caltech.edu/CaltechAUTHORS:20130305-102043001. It seems that the more heterogeneous the slip, the higher the average stress drop.
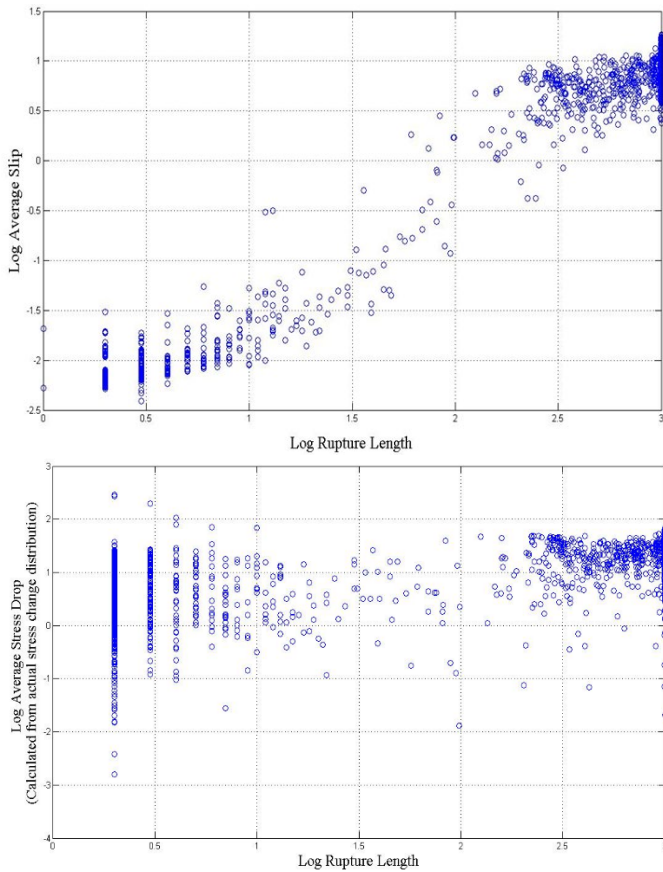
Figure 8.55. CaseA

There are several parameters to play with in the sbm models. The first and most important is the ratio between the coil spring stiffness and the leaf spring stiffness. If the coil springs are very stiff, then the blocks tend to move in unison and then there are more large events and fewer small events (Case A, Fig. 8.53). That is, stiffer coil springs produces events with less spatial heterogeneity and it also results in suites of events that have lower b-values (Gutenberg-Richter).

Figure 8.56 shows the events produced by a system with more compliant coil springs (Case B). Notice that there are more moderate sized events than Case A with stiffer coil springs. That is, this system has a lower b-value. Although the average stress drops for this system are higher than for case A, they are still scale invariant. The scale invariance seems to come from the fractal nature of the prestress. Figure 8.57 shows that the prestress for case B is more heterogeneous than for case A.
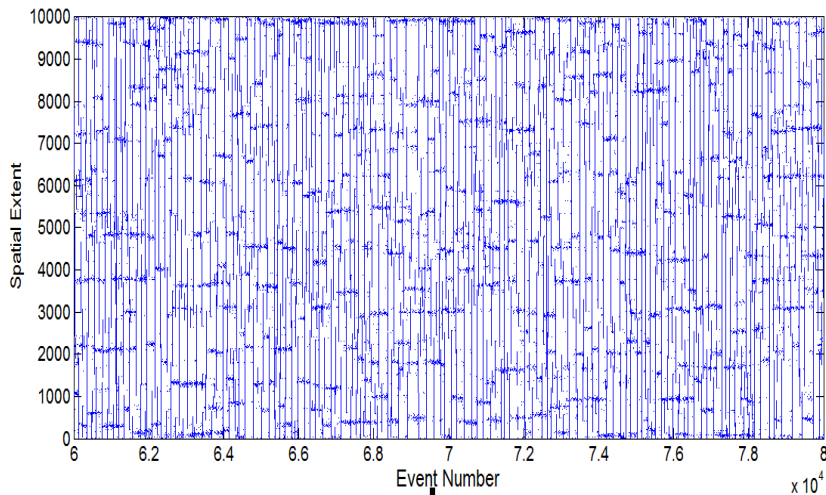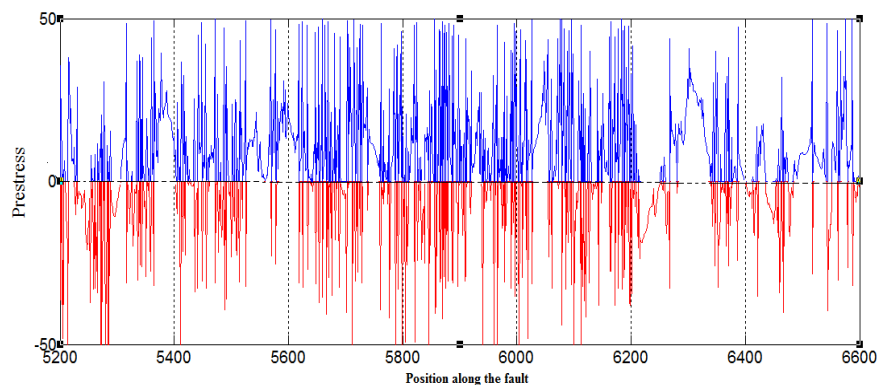
Figure 8.56. CaseB

Fig. 8.57. Case B

How to interpret the chaotic sbm

The sbm system has the advantage that slip-pulse dynamics can be studied with modern computers, whereas 3-d elastic systems cannot, even if the events are confined to a single planar fault. Nevertheless, it's important to acknowledge that there are important differences between the Earth and an sbm. For example, we can tune an sbm to produce power-law frequency-magnitude statistics (similar to the Gutenberg-Richter law). But the sbm confines all of the events to a single line (it's a 1-d model). In contrast, it seems clear that real earthquakes are distributed within a volume. As modern seismology is providing ever greater fidelity to resolve spatial patterns in seismicity, we are seeing ever more complex structures. That is, seismicity seems to not be amorphous clouds of events. Instead, seismicity seems to occur on complex assemblages of planar faults. The implication is that earthquakes occur on fractal networks, while at the same time, the dynamics of the failure is fundamentally chaotic and described by fractals.

At this point, it seems hopeless to attempt to model all this geometric and dynamic fractal complexity. Perhaps a more constructive question is to ask what we would do with our models if we were capable of realistically simulating earthquake phenomena. My opinion is that these new dynamic models provide

us with a deeper understanding of failure processes, and they hint at new frameworks for characterizing mechanical properties.

The Earthquake Cycle

Several decades ago, Keitti Aki proposed that all earthquakes shared simple common characteristics (i.e., stress drop, rupture velocity, geometric aspect ratios). Even though I have argued that his insight was fundamentally flawed, Aki's conjecture has been a useful framework to understand the characteristics of earthquakes. Actually, Aki recognized that spatio-temporal complexity was missing from his conjecture and he later proposed models such as the **barrier model** (patches of high mechanical strength) to introduce complexity. Kanamori and Anderson's seminal paper on earthquake scaling was an extension of Aki's conjecture. Kanamori later introduced the **asperity model** (patches of high prestress) to add spatio-temporal complexity into the model. Although barriers and asperities are fundamentally different, they serve the same purpose of adding complexity.

Whether it's barriers or asperities, both of these models lead to crack-like dynamics; neither of these models are chaotic. Their behavior may be complex, but it is fundamentally predictable; it's ultimately controlled by spatial variations in earth materials that define friction. Based on the behavior of these crack-like models, researchers have proposed that earthquakes are fundamentally dominated by repeating large earthquakes. A cartoon of this view of earthquakes as large events that repeat regularly (controlled by the average slip per event and the fault loading rate) is often referred to as the **seismic cycle** (see Fig. 8.58).

Geologic evidence of slip in past earthquakes has suggested that complex slip patterns repeat from one slip event to the next one. This has motivated the notion that the large earthquakes are **characteristic earthquakes**. I often observe colleagues who opine about where we are in the seismic cycle. For example, "the big one is overdue." While this type of earthquake cycle logic is natural if earthquakes are crack-like ruptures, it seems nonsensical from the viewpoint of chaotic slip pulses. Slip pulses arise from extreme spatio-temporal variations in dynamic friction. Dramatically low dynamic friction only occurs in the core of a propagating slip pulse.
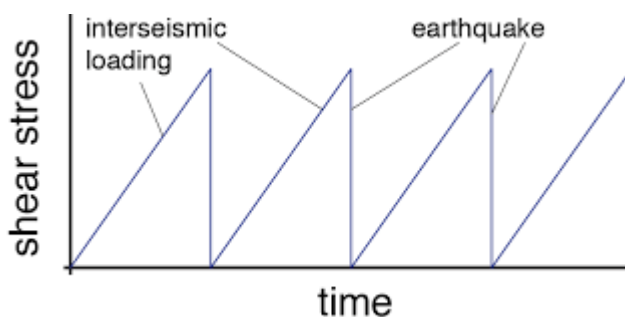

Figure 8.58

One way to get a feel for importance of the crustal strength is to estimate the total strain energy in the Crust of southern California and to then estimate how many earthquakes would be required to remove all of this energy. This is kind of a whimsical calculation, but it can help to build intuition.

I begin by assuming that the shear stress in rocks is close to the limit of Byerly friction (I have heard several colleagues make this claim). That is, I will assume a coefficient of friction of 0.6 and that between the free surface and a depth of 15 km, the shear stress is approximately increasing linearly at a rate of 20 MPa/km. Now the strain energy density is given by $\frac{1}{2}\sigma\varepsilon = \frac{1}{2\mu}\sigma^2$. Therefore,

$$
\begin{aligned}
E &= \iiint \frac{1}{2}\sigma\varepsilon \mathrm{dV} \\
&= \frac{1}{2\mu}\left(\text{Area of s.CA.}\right)\int_{0km}^{15km}\sigma^2\mathrm{d}z \\
&= \frac{1}{2\mu}\left(\text{area of s.Cal.}\right)\int_{0km}^{15km}\left(\frac{z\cdot 20\text{MPa}}{\text{km}}\right)^2\mathrm{d}z \\
&= \frac{4\cdot 10^2\,\text{MPa}^2}{2\cdot 4\cdot 10^4\,\text{MPa}}\left(\text{area of s.Cal.}\right)A\int_{0km}^{15km}\left(\frac{z}{\text{km}}\right)^2\mathrm{d}z \\
&= \frac{\text{MPa}}{2\cdot 10^2}\left(\text{area of s.Cal.}\right)\int_{0km}^{15km}\left(\frac{z}{\text{km}}\right)^2\mathrm{d}z \\
&= \frac{\text{MPa}}{2\cdot 10^2}\left(\text{area of s.Cal.}\right)\frac{1}{3}\frac{(15km)^3}{km^2} \\
&= 10\text{MPa}\cdot km\cdot\left(\text{area of s.Cal.}\right) \\
&= 562\cdot MJ\cdot\left(\text{area of s.Cal.}\right)
\end{aligned}
$$

8.119

Now let's take the approximate area of southern California to be 300 km by 300 km or $10^{11}m^2$. Then

$$E \approx 5\times 10^{20}\,J \tag{8.120}$$

Now the conversion to energy magnitude is

$$M_W = \frac{\log E - 4.8}{1.5} \tag{8.121}$$

So there is enough energy for a M= 10.6. Or alternatively, if the crust is uniformly at the Byerly limit, then there is enough strain energy for 7,500 M 8 earthquakes. In this whimsical world, the next earthquake nucleation could set off a chain reaction of 7,500 M 8 earthquakes. Clearly the assumption that the stress is near the Byerly limit is incorrect.

Strength of the chaotic spring-block slider model

While figures similar to 8.58 are common in earthquake studies, they typically avoid putting numbers on the vertical axis to describe shear stress. Nevertheless, there is an implied maximum shear stress that many researchers describe as the strength of the system. As I have discussed earlier, there are long-standing debates about what is the appropriate strength of the crust. Before, I discuss the strength of the crust, I want to discuss the strength of the sbm just presented. This system provides some important insight into how to view the strength of the crust.

Strength based on average stress

If the prestress in the sbm was uniform, then defining the strength of the system would simply be the stress at which failures occurred. However, the prestress is very complex. Given this complexity, there are several different ways that we could think of to measure the strength of the system.

Let's begin by looking at the prestress in the event shown in Fig. 8.52. We could define the strength to be the maximum shear force that can occur on a block, which in this case, is 50. However, the choice of maximum force of 50 has little effect on the overall behavior of this system. A more meaningful choice for strength would be to measure the total shear force on all the blocks that moved in the event that occurred as a result of the force distribution in Fig. 8.52. In a laboratory experiment to measure strength, we typically measure a load that results in failure, and then we divide by the cross-sectional area of the sample to obtain the strength of the material (described as a stress, or force per area). I will define this to be the average stress-based strength, or

$$\Gamma_{\bar{\sigma}} \equiv F\big/L = \frac{\int_{o}^{L} \sigma(x)\,dx}{L} = \bar{\sigma}_0 \qquad\qquad 8.122$$

In the case of the event for fig. 8.52, the average stress is less than 10, which is much less than the maximum shear stress of 50.
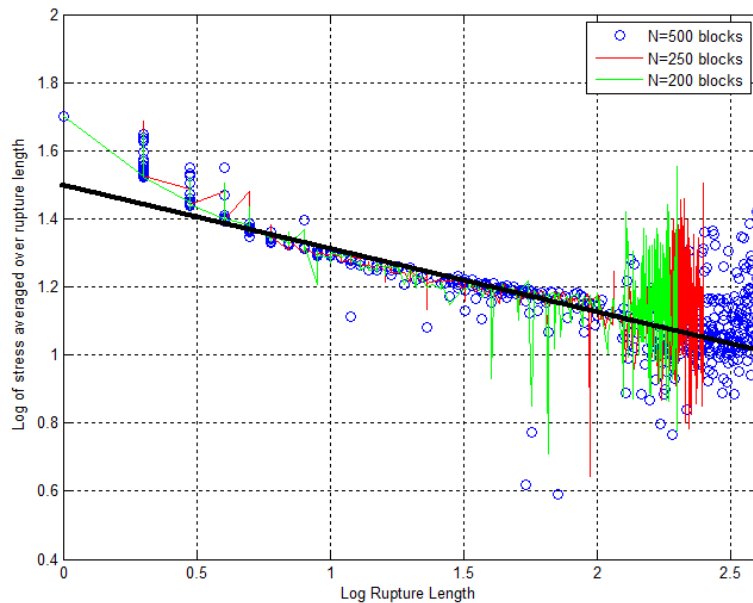
Fig. 8.59.   Case A $\Gamma_{\bar{\sigma}} \approx 32 L_R^{-0.2}$

We can immediately see that if the Earth experiences slip-pulse failure, then the strength of the Crust is significantly lower than the maximum shear stress in the system (presumably the shear stress at the hypocenter).   That is, laboratory experiments on the frictional strength of materials may not tell us much about the strength of the crust. My opinion is that the "Christmas tree" model (see Fig. 8.26) does not represent the strength of the crust. Instead, it provides an estimate of the maximum localized stress.

In the rate-weakening sbm, faster sliding is associated with lower friction. Faster sliding is more common in longer ruptures that have larger slip pulses. These physics are behind the correlation between strength of the sbm and the length of the ruptures. That is, the average stress decreases with the rupture length as is shown in Fig. 8.59. The heavy line is a log-log fit that indicates that average stress based strength scales with rupture length as, $\Gamma_{\bar{\sigma}} \sim L_R^{-0.2}$. Obviously, larger dimension solids (e.g., the Crust) support larger events; the larger the event, the lower the strength.

This may, at first, seem counter-intuitive. Perhaps it's easier to see if you realize that the definition of strength that I am using involves finding the average of the self-organized prestress at different lengths. It seems that the longer the averaging length of prestress, the lower the average. I will get back to that shortly when I discuss power laws.

Strength based on work

In the laboratory, it is straightforward to calculate the average stress in a sample. However, this is not a feasible measurement in the Earth since we don't know the total load on the failure surface. In principle, we could measure the stress tensor at enough locations that we could obtain a realistic average. In fact, that is what would be required if we want to take a spatial integral to estimate total load. If

stress is heterogeneous, then it is not feasible to measure it at enough points to spatially integrate it.

Perhaps you have already observed that I have previously discussed the strength of the crust using arguments about heat energy, fracture energy, and radiated energy (see eq. 8.89). In the sbm, there is no radiated energy or fracture energy, frictional heating is the only dissipation to absorb the potential energy released by the springs. Accordingly, I define **strength based on work** as

$$\Gamma_W \equiv \Delta E_{dissipation} \Big/ L\bar{D} \qquad\qquad 8.123$$

You can think of Γ_W as being analogous to a plastic yield stress. It provides an estimate of how much work is required to change the material from an initial configuration to a final configuration. Since energy is an integrated quantity, it is not necessary to worry about the spatial irregularities associated with fractal stress.

Figure 8.60 shows the length scale dependence of Γ_W, and it should be compared with Fig. 8.59. There is a considerably stronger weakening of this strength with increasing length scale. In this example $\Gamma_W \approx 63 L_R^{-0.46}$ (it's case A, see Figures 8.54 and 8.55). The fact that longer ruptures are associated with higher slip velocities is the reason for this stronger length dependence of the energy based strength.
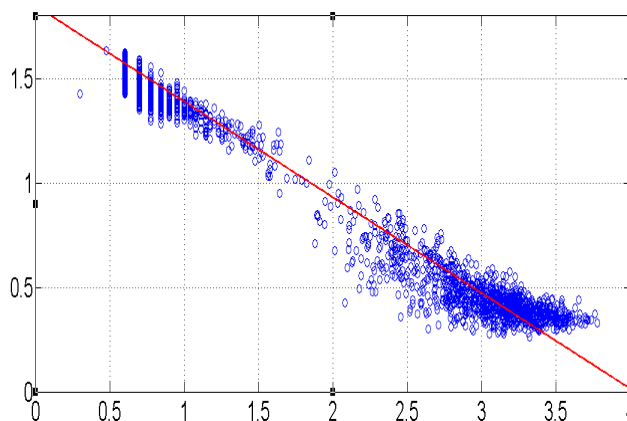
Figure 8.60. log-log Plot of energy-based strength, Γ_W, vs rupture length for events from case A of the sbm. Notice that there is a stronger dependence of energy-based strength on rupture length than there is for stress-based strength (see Fig. 8.59).

This is an important lesson from the sbm. That is, **when considering materials that experience slip pulses, determining the strength of the system using analyses based on energy considerations is likely to provide smaller estimates than is obtained by estimating the total load on the region that fails**.

Figure 8.61. Case B

$$\Gamma_W \approx 10 L_R^{-0.19}$$

Power-Law Scaling

While crack-like rupture dynamics can be quite complex, the far-field radiation from these models can be described by some scaling relations about amplitude and duration. Fundamental theorems concerning the principle of "equivalent width" can be used to infer the shape of Fourier amplitude spectra for crack-like models. Crack-like models with constant stress drop typically produce amplitude spectra that can be characterized with simple power laws. In particular, these power laws have simple ex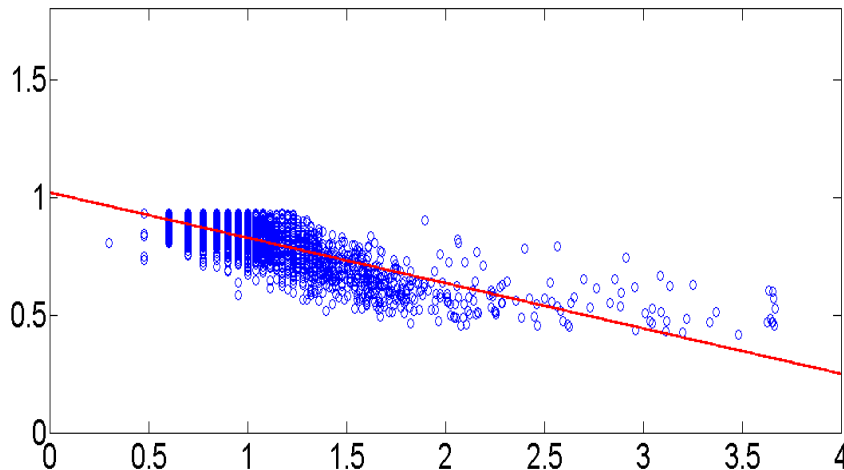ponents (e.g., $f^{-2}, f^{-3/2}$). These laws have exponents that are rational numbers because of some simple geometries (e.g., rupture area scales with L^2, stress drop scales with D/L, derivatives and integrals are equivalent to multiplying or dividing f or f^{-1}, etc.).

However, in chaotic self-organized systems, there are also typically power laws, but in this case, the exponents can be any real number (Although this is a speculation, I suspect that these exponents are typically an irrational number). The power law behavior is the result of self-organization of the system into a self-similar fractal. By definition, fractals are scale free.

I believe that self-organization of chaotic dynamics that is the key to understanding Aki's observation that earthquakes seem similar over a wide band of length scales. Ironically, it became popular to erroneously use the term self-similar for Aki's conjecture (his model is similar, but not self similar). But in the end, the more appropriate model of earthquakes as scale free self-organization of chaotic dynamics produces true self-similarity.

It may seem that we are at an impasse. Models like the sbm provide the insight that we are dealing with a chaotic self-organizing system, but computational limitations prevent us from meaningful simulations of 3-d solids that yield during strong rate-weakening friction.

One approach to this impasse is to develop new equations that provide us with meaningful calculations, while at the same time requiring fewer calculations. The need to allow the system to self-organize while simultaneously honoring the complexity that comes from positive feedback dynamics is why this is so difficult. It's important to recognize that if the solution is described by a fractal, then there is little hope of using conventional continuum mechanics (Navier's equation). That is, Navier's equation is a 2nd order partial differential equation, whereas our solutions likely look like fractals. A fundamental difficulty with fractals is that the conventional definition of a derivative doesn't make sense for fractals. That is, derivatives are defined to be the change in a function's value as the step size shrinks to zero. Fractals are irregular at all scales, even when the step size shrinks to zero. This means that **this class of problem can only be investigated with discrete models**. In effect, we anticipate that we will be unable to derive exact solutions to problems.

It is important to recognize that the Fourier spectra of parameters describing individual events that occur as a result of self-organized chaotic dynamics do not necessarily follow a strict power law. Instead, it is **the expected value of** these parameters that **obeys a power law**.

Although I have argued that the amplitude spectrum has a power-law expected value, I have avoided discussion of the phase spectrum. This is a tricky issue, since chaotic behavior is definitely not the same as random phenomena. That is, there is structure (often hidden) that results from these systems. This structure is fundamental to the identification of **strange attractors.** In the case of our problem, we can anticipate that the phenomena are **statistically stationary**. That is, the expected values of our parameters are independent of the origin system.

Slip v Length (stress drop) scaling

One example of a stochastic power-law model comes from a study of slip vs length scaling by Liu-Zheng, Heaton, and DiCaprio. We considered a simple 1-d model of slip as a function of linear distance x. In particular, we assumed that slip is a stochastic function of x. We hypothesized that the amplitude of a slip pulse as it propagates along the fault is described as a **random walk**. Consider that I have discretized distance along the rupture as x_i. I will assume that

$$D(x_{i+1}) = D(x_i) + R_i \qquad\qquad 8.124$$

Where R_i is a number chosen from a bin of random numbers. This is known as a 1-d random walk and this process has been used to simulate the **Brownian motion** of particles that are randomly impacted with fluid molecules. 1-d Brownian motion is the same as taking the integral of random white noise over all past steps. The random sequence $R(x)$ has a Fourier transform that is also composed of random numbers. That is, the spectrum looks like random numbers that do not change their expected values as a function of wavenumber. Therefore, the amplitude spectrum is flat and this is called **white noise**. In

contrast, the stochastic slip is the spatial integral of this random function and thus it has a $\frac{1}{k}$ spectrum.

I can generalize this model by assuming that the slip is a fractal function of space, or

$$\tilde{D}(k) = \tilde{D}_0 \frac{R(k)}{k^\alpha} \qquad\qquad 8.125$$

In standard Brownian motion, $\alpha = 1$. However, since we are dealing with self-organizing fractals, we anticipate that α may be an irrational number; in this case the process is called a fractional random walk. Unfortunately, current measurements of slip variations are inadequate to obtain a stable estimate for the amplitude spectrum of slip of real earthquakes.

Liu-Zheng, Heaton, and DiCaprio (2005, GJI, doi:10.1111/j.1365-246X.2005.02679.x) used this fractional random walk model to reproduce measured values of the ratio of average slip divided by the rupture length. In particular, measurement of Potency and rupture length are available from many earthquakes. As I will now show, the ratio of Potency to rupture length depends on the value of α.

We considered this to be an example of a class of problems known as **gambler's ruin**. The focus of this problem is to estimate how many games a gambler can play before he loses all of his money. That is, slip as a function of space is transformed into a sequence of games of chance. Once the slip crosses the x-axis, the earthquake is considered to be over. That is, the rule of this game is that earthquakes are comprised of spatially continuous regions of slip. Figure 8.62 shows examples of spatially contiguous slip profiles for different values of α. All of the walks start at zero slip and end at zero slip; most of the examples involve only several steps until the slip returns to zero. The three examples in Fig. 8-62 are chosen to show the spatial characteristics of the larger events in this numerical experiment. The spatially averaged slip for each event is shown as D1, D2, and D3. Notice that the red curve $(\alpha = 1.0)$ is quite rough. This is true Brownian motion. The length of this red rupture is 72 km and the average slip is 4.5 m, which gives a slip to length ratio of 6×10^{-5}, which is comparable to the average value of earthquakes (see 8.33).

The next thing to notice in Figure 8-62 is although the different events have comparable rupture lengths, the average slip decreases as the roughness is filtered out. This trait is clearly displayed in Figure 8-63, which shows how increasing smoothness, α, results in systematically longer rupture for a given average slip. That is, the rougher the fractional random walk, the better the chance that the slip will cross the x-axis (termination of the event).

You should recognize that the ratio of average slip to rupture length is a measure of the average stress drop for these events. Thus, this model predicts that **rougher slip distributions produce higher average stress drops**. Also notice

that, although the red curve $(\alpha = 1.0)$ has a rupture length of 73 km, it very nearly terminated at the 67-km mark, which would have been a 44-km rupture. That is, it was mere chance that the red event did not have a significantly higher stress drop. In this model, two events could have identical slip distributions up until one of them stopped rupturing. If the second event continued to rupture it, might end up with a much lower stress drop, even though the radiation from the first part of the ruptures was identical. This **example shows that average stress drop may not tell us much about the dynamics** of a rupture.
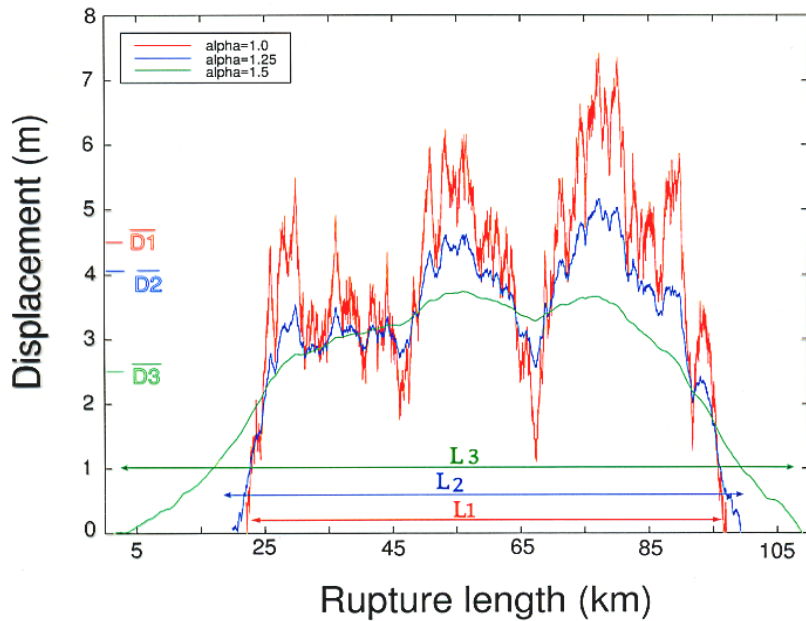
Figure 8-62
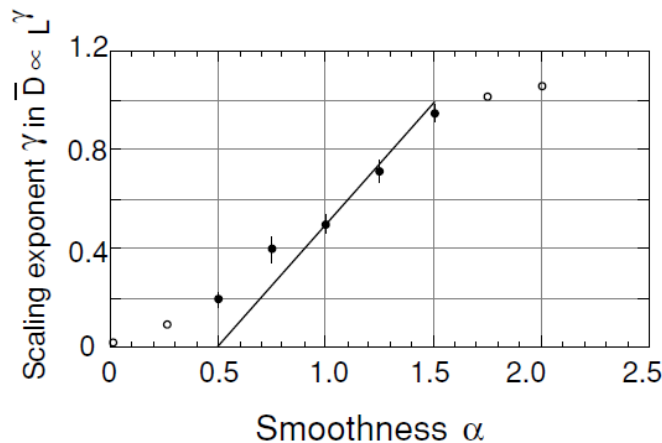
Figure 8-63.

Now the observation that average stress drop is approximately independent of scale is direct evidence that average slip scales linearly with rupture dimension. Notice that in standard 1-d random walk (aka, Brownian motion $\alpha = 1.0$), the expected rupture length scales as the square root of the average slip. However, when the smoothness increases to $\alpha = 1.5$ (one and one half integrals of random white noise), then the expected rupture length scales linearly with average slip.

8-107

Curiously, $\alpha = 1.5$ seems to be some type of critical point. That is, if $\alpha > 1.5$, then the random walks tend to diverge with distance (the gambler always wins).

While a 1-d fractional random walk model can adequately explain stress drop scale invariance, it does a poor job of simulating the Gutenberg-Richter relationship. Figure 8-64 shows the slip/length ratios of many simulated events And for several values of σ. It appears that a smooth distribution , $\alpha = 1.5$, produces just as many long ruptures as it does short ones. That is the b-value is approaching 0. In contrast, a standard random walk ($\alpha = 1.0$) produces many more short ruptures than long ones.

Figure 8-64

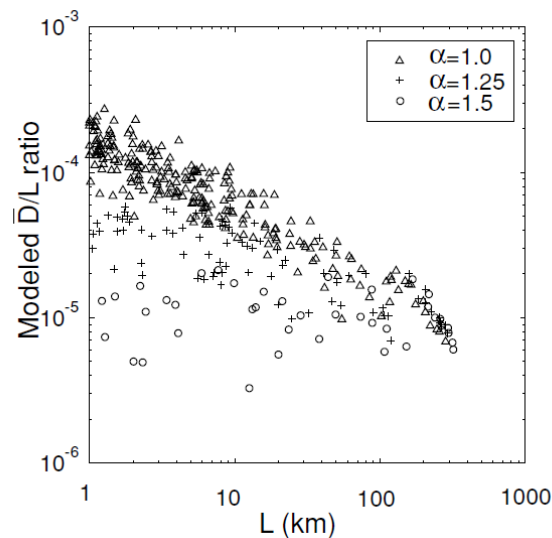This observation raises some interesting questions. Are large strike-slip ruptures with very long ruptures (340 km for 1857 Ft Tejon, or 450 km for 1906 San Francisco) compatible with the Gutenberg-Richter law? If they are to be explained by a fractional random walk, then the slip must vary slowly along the rupture. I admit that I don't know how to explain these very long ruptures. However, if you examine the slip to length ratio of major historic crustal earthquakes (see Wells and Coppersmith, BSSA, 84, 974-1002), you will see the distribution shown in Fig. 8-65a. In comparison, Fig. 8-65b shows the result of a combination of numerical simulations of fractional random walks with $\alpha = 1.25$ and $\alpha = 1.5$. This particular combination matches the observation that the variation in the slip to length ratio is larger for shorter ruptures than it is for longer ones (notice the trend in the upper limit in both plots).

The observed similarity between measure slip/length ratios and fractional 1-d random walks is evidence that the observed scale invariance of stress drop may be closely related to the heterogeneity of ruptures, and not a simple characteristic of friction. The fact that, for a given magnitude, stress drops can vary by more than two orders of magnitude has often been attributed to modeling error. Personally, I prefer to interpret this large range of reported stress drops as being a real phenomenon that just reflects variations in random walks.
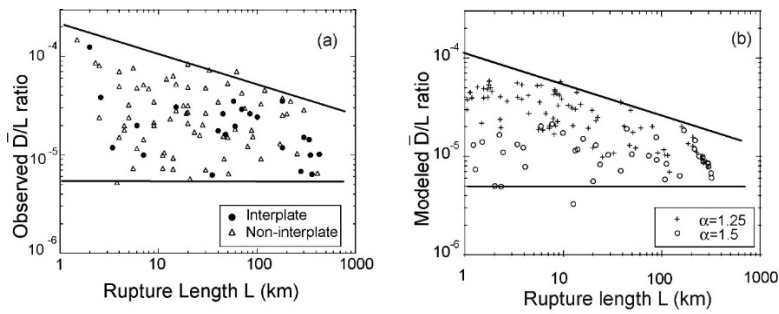
Figure 8-65

It is important to recognize that the models of spatial complexity that I have just described all come from 1-d models; there are no complicated fault systems similar to what is observed in the real earth. In fact, there is great variety in the geometric complexity of fault systems. Some faults are remarkably simple planar unconformities. In fact, it is possible to view the Punchbowl fault in great detail by a pleasant trip to Punchbowl Regional Park. The Punchbowl fault is a major right-lateral strike-slip fault that runs sub-parallel to the San Andreas fault just north of the San Gabriel Mtns. This fault is interpreted to be an early version of the San Andreas and it appears that the total offset on this fault is about 45 km. Luckily, erosion over millions of years has exposed extensive sections of the fault for detailed investigation. There is a very good field guide by Fred and Judy Chester (they have many excellent publications describing many aspects of this fault). A visit to the Punchbowl allows you to view the fault along its strike and it's plain to see that it is geometrically flat (despite the large total offset). It's even more amazing to view the fault up close. In most areas, the total thickness of the fault zone is less than a meter and there is evidence that most of the sliding was confined to a zone only several mm wide.

The San Gabriel Fault is another early version of the San Andreas that has been exposed by erosion (again studied by Chester and Chester). If you hike along this fault zone, you will find some places where the fault is a simple thin zone (like the Punchbowl fault), but you will also find other places where the fault becomes very complex (a complex zone of pulverized rock more than several hundred meters wide (for example, visit Red Box Gap just north of Mt. Wilson).

Recent advances in remote imaging have provided a new perspective on the complexity of individual earthquake ruptures. For example, comparison of before and after satellite-based photographs of the desert floor show that the 2019 Ridgecrest rupture occurred on a complex network of fault structures (Figure 8-66). Given the observed geometric complexity of observed faults, it's easy to think that the dynamic complexity of earthquakes is entirely due to this geometric complexity. This is an important unsolved problem in earthquake physics. That is, we suspect that 3-d complexity of faults is important, but we can't even numerically simulate the dynamic complexity that occurs on simple planar faults that experience pulse-like ruptures.
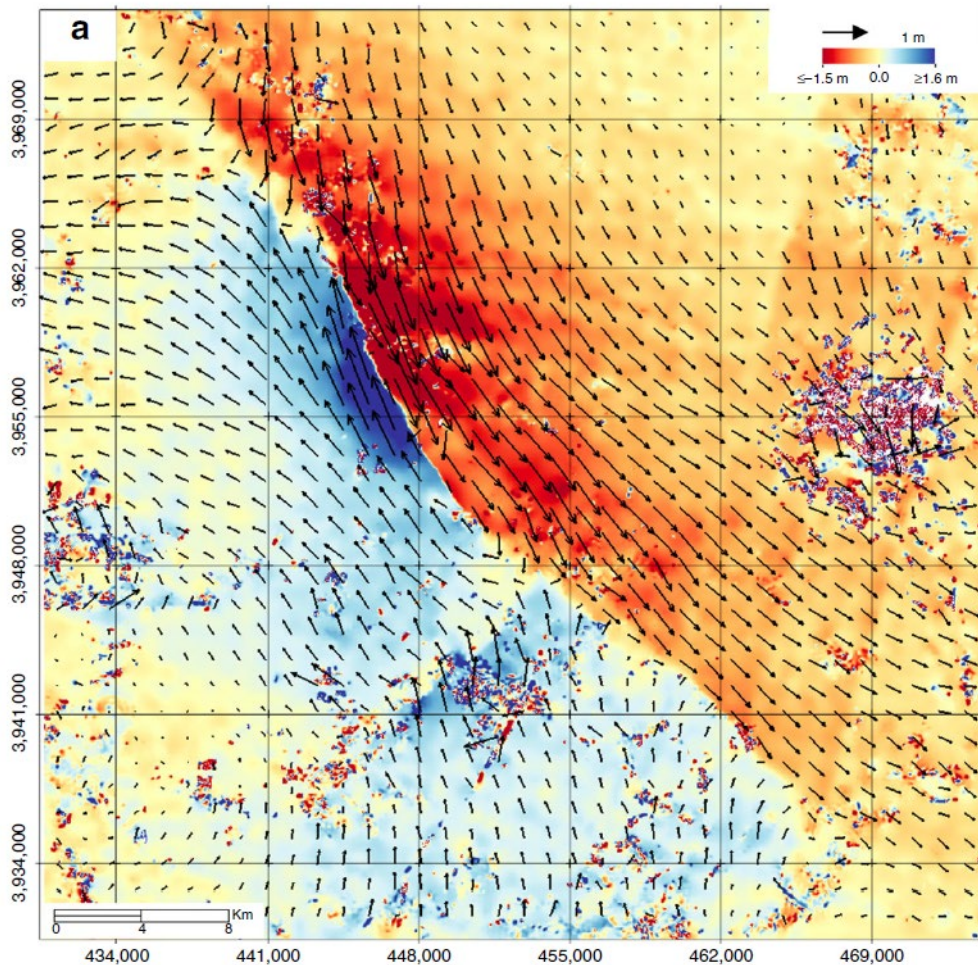
Figure 8-66. Deformation of the ground surface in the vicinity of the 2019 Ridgecrest earthquake sequence. (Chen, Avouac, Aati, Milliner, Sheng, and Shi, https://doi.org/10.1038/s41467-019-13750-w).

The type of complexity shown in Fig. 8-66 is typical of other well-observed crustal ruptures (e.g., 1991 Landers, 2001 Hector, 2011 Cucapah-El Major, 2017 Kumamoto, 2018, and 2016 Kaikoura). In contrast, there are long sections of the San Andreas fault that appear to be relatively simple and planar (especially the Carizzo Plain). The very long ruptures in 1857 and 1906 appear to have been on relatively straight sections of the San Andreas.

Although we observe geometric complexity in fault systems, it is difficult to simulate and interpret. In particular, any kind steadily rupturing fault that has localized jumps in fault orientation (e.g., a kink) will produce singular stress changes at geometric corners. Some modelers have argued that geometric corners are the likely places to initiate and terminate rupture (often referred to as **fault segmentation**). Although this seems like a plausible assumption, there are clear examples where fault ruptures seem to ignore these segment boundaries.
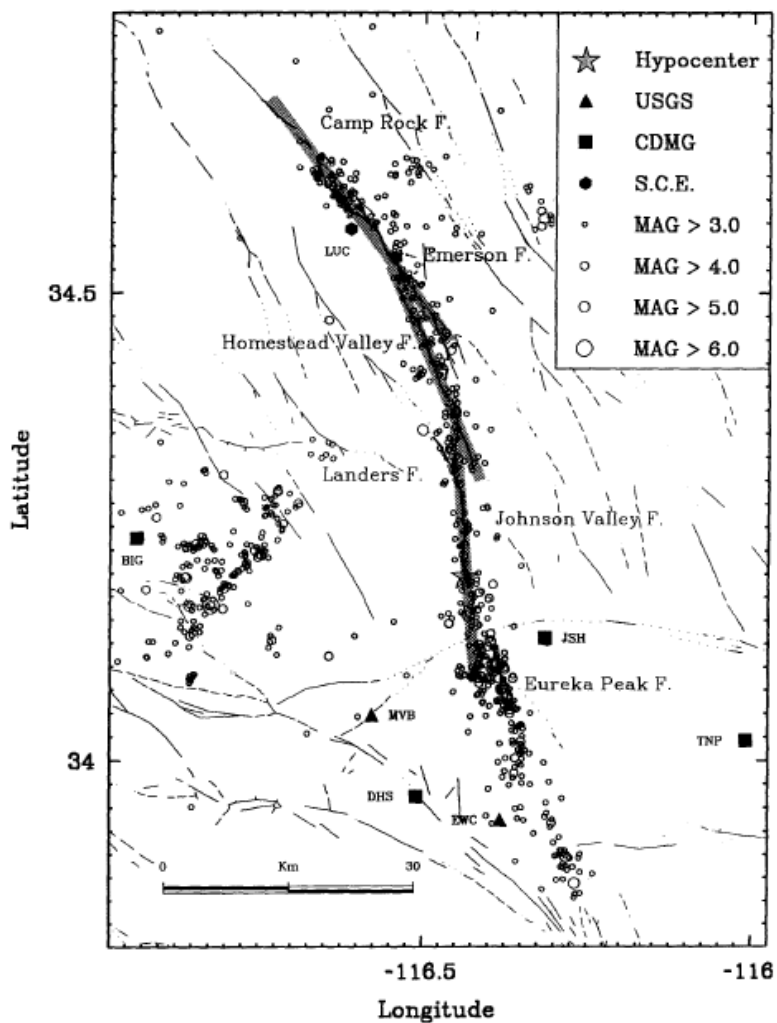
Figure 8-67. Geometric complexity of the 1991 Landers earthquake.

The 1991 Landers rupture is an example that shows a failure of the fault segmentation hypothesis. That is the rupture progressed northward along the Johnson Valley Fault, and then it continued to rupture when it intersected the Homestead Valley Fault. It then ruptured right through another segment boundary between the Homestead Valley Fault and Emerson/Camp Rock fault. The rupture finally terminated in the middle of a fault segment (perhaps the straightest fault segment in the whole complex). While it may be that fault complexity and segmentation has a first-order effect on rupture dynamics, the Landers earthquake clearly shows that simple models do a poor job of predicting reality.

Heterogeneous Stress in the shallow Crust

There are numerous places in this chapter that I have argued that stress is likely to be spatially complex (look at the discussion of chaotic slip pules). In the following, I discuss the statistical characteristics of this heterogeneous stress. Unfortunately, this discussion may be difficult to follow. Furthermore, the hypothesized model is very different from conventional models of crustal stress (for example, see Zoback and Zoback, or Scholz). The model was developed by Deborah Smith and me and it is designed to show that the effect of heterogeneous

stress depends strongly on the length scale that is appropriate for different problems.

Figure 8-68. Map of Southern California with velocity arrows used in this study in a fixed North America reference frame. Red arrows: velocity provided by the Scripps Orbit and Permanent Array Center (SOPAC); blue arrows: data from the University of Miami; green arrows: data from the cru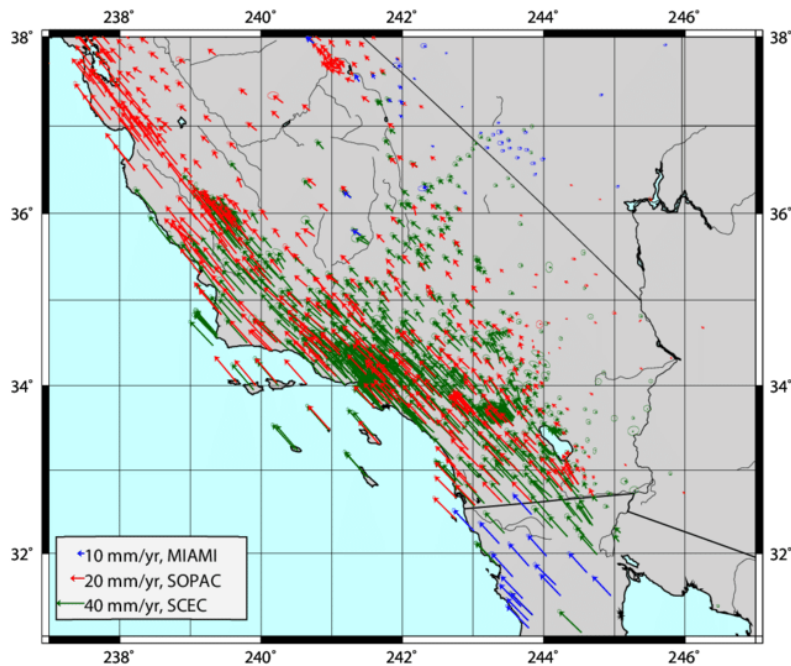stal motion map (version 3) by the Southern California Earthquake Center (SCEC). Hackl, M. & Malservisi, Rocco & Wdowinski, Shimon. 2009, Strain patterns from dense GPS networks, Natural Hazards and Earth System Sciences, 9. 10.5194/nhess-9-1177-2009.

The measurement of stress in the Crust is actually very difficult. Testing in deep boreholes is probably the most direct approach to this problem. Unfortunately, deep boreholes are extraordinarily expensive and they are typically limited to the upper 5 km of the Crust. Since direct measurements are so sparse, many earth scientists use simple models to infer the stress in the crust. One of the most successful methodologies to inform these models comes from observations of the spatio-temporal distribution of the changes in positions of geodetic monuments over time. The introduction of satellite-based surveying (Global Positioning Satellite system, GPS) in the 1990's has provided the data to track the steady deformation that drives plate tectonics. Figure 8-68 shows the velocities of stations in southern California relative to the average velocity of North America. These velocity vectors are very coherent; they all point to the northwest and their amplitudes steadily increase as for stations located further from stable North America.

The impressively coherent, long-term, point velocities in Fig. 8-68 might mislead you into the impression that the crustal stress in comparably coherent. In contrast, Fig. 8-69 shows the orientation of the principal compression axes measured in a number of boreholes in southern California. These orientations are primarily obtained by observing the orientation of **borehole breakouts**, which

are spontaneous fractures in the walls of boreholes (primarily from the oil industry). Borehole breakouts typically occur in the faces of the borehole that are perpendicular to the least compressive principle stress axis (aka, the Sh axis). SH is the direction of the largest compressive principal stress axis. B
because the Earth's surface is traction free, Sh and SH are generally horizontal and therefore, they are perpendicular.
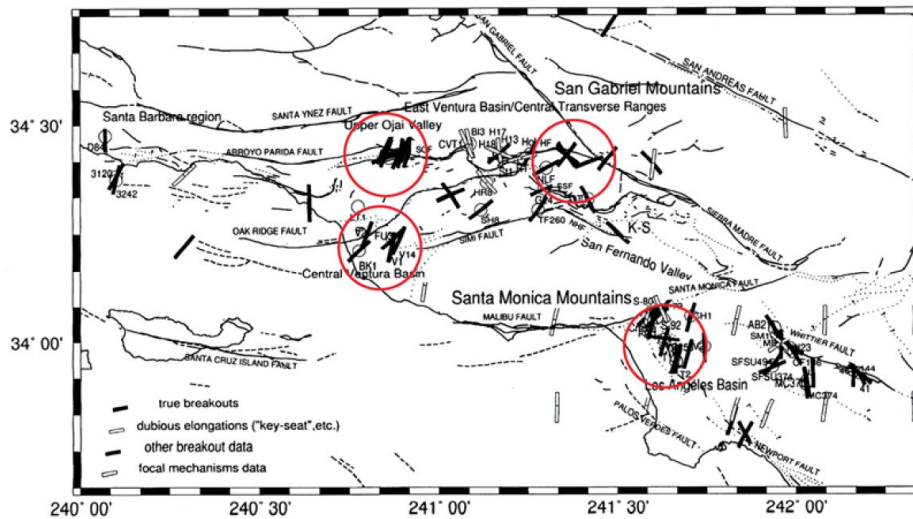
Figure 8-69. Wilde and Stock [1997, *J.Geophys. Res.*, *102*, 4969-4983] plotted inferred maximum horizontal compressive stress, SH , orientations from borehole breakouts in Southern California. There are a variety of orientations for borehole breakouts from the same borehole or from boreholesspatially close to one another. This suggests short-wavelength spatial stress heterogeneity. In this modified plot, we have used circles to point out a few of the locations studied by Wilde and Stock that show evidence for SH orientation heterogeneity. Modified by Smith and Heaton, 2011, Bull. Seism. Soc. Am., 101, 1396–1421, doi: 10.1785/0120100058

Notice that there is considerable scatter in the SH orientations. Although this scatter is sometimes described as "noise," observation of breakouts is relatively straightforward. That is, direct measurements of stress orientation provide evidence that the actual stress field may be far more complex than you might imagine from inspection of particle velocity field shown in Fig. 8-68. It's not noise.

The statistical analysis of the orientations of focal mechanisms is currently the most popular method to determine the spatial orientation of stress in the crust (see foundational papers by Angelier, J., 1984, J. Geophys. Res., 89, 5835–5848, and by Michael, A., ,1984, J. Geophys. Res., 89, 11517–11526 and also Michael, A., 1987, J. Geophys. Res., 92, 357–368. These studies make the following **key assumptions**: 1) stress is approximately uniform in space, 2) the crust has a large suite of pre-existing fault planes with a large variety of orientations, 3) these faults have a large variety of yield stresses, 4) seismicity in a region occurs on a diverse suit of planes with many orientations 5) (most importantly) the slip vectors of all of the earthquakes are aligned with the direction of maximum shear

stress in the region.   According to this technique, the orientation of the principle stresses in a region is that orientation that minimizes the misfit of one of two plausible slip vectors (there are two conjugate planes for every focal mechanism) for the suite of focal mechanisms.

Focal-mechanism-based stress inversions are widely used as the basis to describe the stress distribution in the Earth's Crust. Unfortunately, there is a **fundamental inconsistency in the underlying assumptions**. Specifically, it seems impossible to have spatially uniform stress (assumption 1) if there is an assortment of randomly oriented fault planes with a very wide range of yield strengths (see Rivera and Kanamori, 2002, Geophys. Res. Lett. 29, art. no. 1088).

A particularly egregious violation of the assumption of homogeneous stress comes from a study of the aftershocks of the 1989 M 6.9 Loma Prieto earthquake by Zoback and Beroza (1993, Evidence for near-frictionless faulting in the 1989 M 6.9 Loma Prieta, California, earthquake and its aftershocks, *Geology*, *21*, 181-185. Figure 8-70 shows the focal mechanisms of aftershocks of the 1989 Loma Prieta earthquake projected onto the main fault plane. Slip in this earthquake was oblique (a steeply dipping plane with a slip vector that was primarily right-lateral strike slip with some thrusting).   Rather mysteriously, a significant percentage of the aftershocks had slip vectors in the opposite direction to the mainshock. That is, there were left-lateral aftershocks on the obviously right-lateral San Andreas fault. Zoback and Beroza explained this by hypothesizing that the static coefficient of friction was nearly zero on the San Andreas. This would have made the maximum principal stress axis nearly perpendicular to the San Andreas. Zoback and Beroza argued that there were small variations in fault strike and that the shear stress resolved to different polarities as the strike varied.

I find this explanation to be implausible for several reasons. 1) the amplitude of the slip (see the shaded contours) varies by more than a meter within several km. This implies large local strain changes, which implies large local stress changes. This observation seems incompatible with an assumption of uniformly low static stress on the mainshock fault plane. 2) Such a small static friction (less than 0.1) has never been observed in testing and it seems incompatible with the shear stresses that are necessary to support the gravitational load of the Santa Cruz mountains (directly above the rupture).
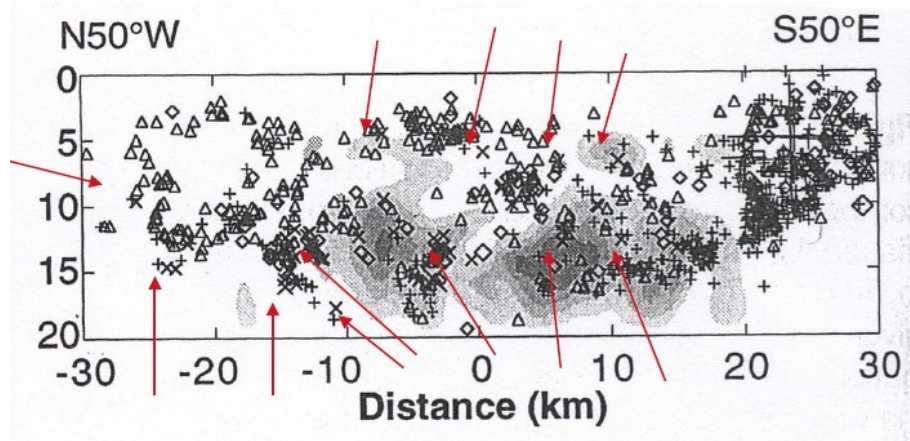
Figure 8-70. Map view of the aftershocks of the 1989 Loma Prieta earthquake (from Beroza and Zoback). The slip distribution derived from a finite-fault slip inversion is also shown as the shaded contours. The vertical axis is depth. (+) represents right-lateral focal mechanisms, (x) represents left-lateral, (Δ) reverse thrusting, and (◊) is normal. Each focal mechanism has two conjugate (perpendicular) planes and the plane that is most parallel with the mainshock is displayed. The Red Arrows (x) highlight aftershocks with slip vectors that are in the opposite direction from the mainshock.

Another explanation is that the stress is highly heterogeneous (similar to the stress in Fig. 8-52) and that there are actually patches of large left-lateral stress that exist on the San Andreas fault. These patches of negative shear stress would be the remnant of previous chaotic ruptures; the dynamic shear stress in the vicinity of a slip pulse are more than an order of magnitude greater than the average stress change. Since steadily increasing tectonic shear stress is in the direction of plate motions, regions of negative shear stress will not nucleate earthquakes. That is, only regions in which the local shear stress is large and also aligned with tectonic stress rates will have future hypocenters. The only time that we expect earthquakes in the opposite direction from plate motion is during an aftershock sequence. That is, the spatially complex slip in the mainshock means that the stress changes from the mainshock are even more complex than slip pattern. This interpretation seems to explain seismicity near the Loma Prieta earthquake; events with negative slips have not been observed in the background seismicity, they are only observed in the aftershocks.

Another curious feature of seismicity (either background or aftershocks) is that it ends to be clustered in space; I sometimes call it clumpy. That is, there are typically spatial knots of persistently high seismicity (just look at any high-resolution seismicity map). Although it may be tempting to assign different material properties to these spatial knots, I suspect that these high-seismicity knots are ephemeral (perhaps on the time scale of decades to centuries). In particular, I suspect that patches of high seismicity are patches of high stress

Deborah Smith and I created a stochastic model of the stress tensor in an attempt to simulate patchy seismicity that has the statistical characteristics of observed focal mechanism catalogs. We hypothesized that, within some region, the

deviatoric stress tensor, $\boldsymbol{\sigma}'(\mathbf{x}, t)$, can be approximately separated into tectonic term that is steadily increasing and approximately spatially uniform, $\dot{\boldsymbol{\sigma}}'^{T}(t - t_0)$, and a term that is spatially heterogeneous and independent of time, $\boldsymbol{\sigma}'^{H}(\mathbf{x})$. This model assumes that at the time of the last nearby earthquake, t_0, the heterogeneous stress and the average background stress, $\boldsymbol{\sigma}'^{B}$, were reset to new values by the heterogeneous stress changes caused by chaotic rupture. The deviatoric stress in a region is written as

$$\boldsymbol{\sigma}'(\mathbf{x}, t) = \boldsymbol{\sigma}'^{B} + \dot{\boldsymbol{\sigma}}'^{T}(t - t_0) + \boldsymbol{\sigma}'^{H}(\mathbf{x}) \qquad 8.126$$

We next hypothesize that events nucleate in the region whenever the shear stress exceeds some pre-determined value, $I_2'^{Y}$, where I_2' is the 2nd deviatoric stress invariant. This failure criterion is referred to as **Henky-Mises plastic yield criterion** and it is commonly used in the ductile yielding of steel. That is, we assume that the shear stress in the crust cannot exceed this limit; once the yield stress is achieved, the material experiences inelastic shearing along conjugate planes of maximum shear stress.

To construct this model, we begin with a 3-d Cartesian grid of discrete points, $x_{i,j,k}$. At each grid point, we use a Gaussian, mean-zero, random number generator to select the values of the components of the stress tensor. We assume that the stress tensor is symmetric, $\sigma_{ij} = \sigma_{ji}$, which ensures that angular momentum is conserved (see Chapter 3). The standard deviation of the diagonal components is chosen to be 1.0, while the off diagonal terms have a standard deviation of $1/\sqrt{2}$. This choice ensures that the **random stress is isotropic**; this is the only ratio that provides stresses that are uniformly distributed in orientation space.

Once we have assigned random numbers to each of the six independent stress components, we apply a **power-law spatial filter** to each stress component. Because our model is isotropic, we can uniquely specify this filter by defining the spectral properties along any line, $r_i = |\mathbf{x}_i|$, passing through the grid. In order to construct the heterogeneous stress, we take the 3-d Fourier transform of each tensor component. We then multiply these spectra by

$$F(k_r) = \left(1 + |k_r|\right)^{-\alpha} \qquad 8.127$$

After which, we take the inverse 2-d transform to obtain the stochastic, power-law, stress tensor in the spatial domain. This process creates a fractal-like distribution. While the expected value of the average components is zero, the mean of a finite number of random numbers is not zero. We remove the mean of our random numbers and then we reintroduce finite mean to our stress through the specification of $\boldsymbol{\sigma}'^{B}$. That is, the stress is the sum of a zero-mean stochastic part with a specified "background" stress.

Figure 8-71 shows the stochastic model of heterogeneous stresses projected onto a plane that intersects the 3-d grid (100x100x100 points).
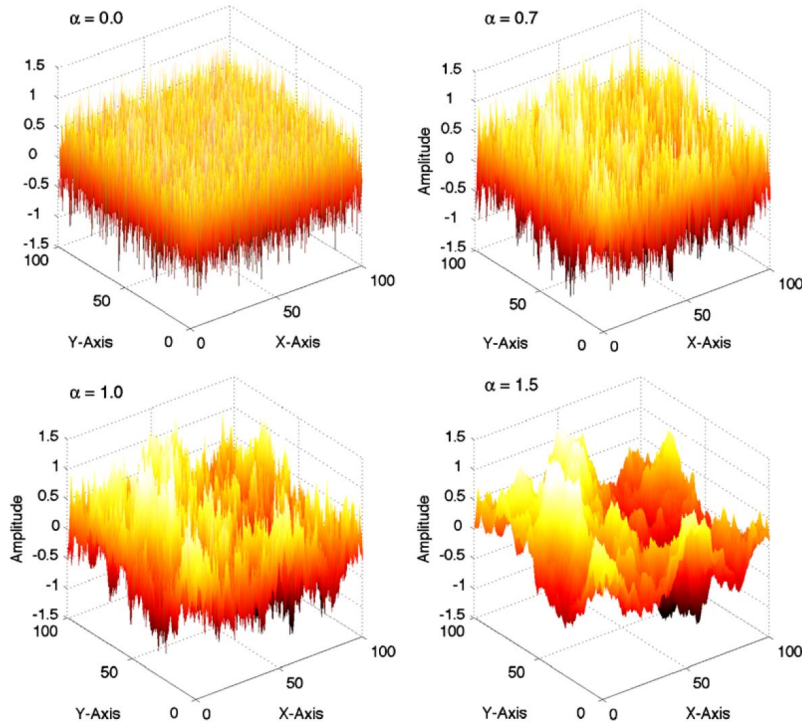
Figure 8-71. Examples of one component of the stochastic power-law heterogeneous stress for different powers, $\left(\dfrac{1}{1+k_r} \right)^{\alpha}$, of wavenumber.

In order to have a 3-d model with realistic values of stress, it is necessary to find a meaningful scale. In particular, we chose to relate the variance of the heterogeneous stress to the size of the uniform background stress σ'^{B}. In particular, we define the **Heterogeneity Ratio, HR**, as

$$HR \equiv \sqrt{\frac{\overline{I'^{H}_{2}}}{I'^{B}_{2}}} \qquad\qquad 8.128$$

Where $\overline{I'^{H}_{2}}$ is the spatial average of the 2nd invariant of the heterogeneous deviatoric stress. $\overline{I'^{H}_{2}}$ is also a measure of the shear-strain energy density (see Housner and Vreeland, 1965). As it turns out, $\overline{I'^{H}_{2}}$ is also the sum of the variances of the components of $\sigma'\left(\mathbf{x}, t = t_0\right)$ (see Smith and Heaton).

For this parameterization of the relative size of the uniform background stress and the stochastic heterogeneous stress, we want to ensure that the modeled stress is independent of the number of grid points used to describe a region. To

accomplish this, we use the same outer scale for different grids, independent of the number of grid points. Fig. 8-72 shows examples of the amplitudes of maximum shear stress (heterogeneous plus background) at points along a line intersecting the grid. Random white Gaussian noise is used for the left column and the preferred power-law, $\alpha = 0.8$, is used in the right column. The heterogeneity ratio, HR, increases from top to bottom; the model in the lower right (HR=2.375 and $\alpha = 0.8$) best fits southern California focal mechanism data.
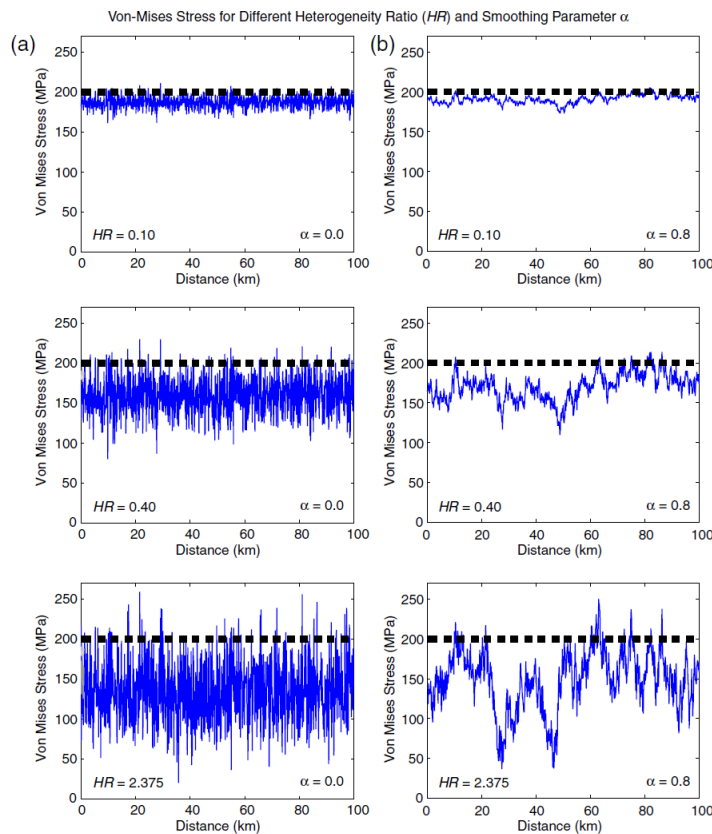
Figure 8.72. Maximum shear stress (aka, Von Mises stress) plotted along lines that intersect the 3-d spatial grid of fractal stress. The horizontal dotted lines are taken to be the local yield stress of the material (200 MPa). The overall heterogeneity increases from top to bottom. The heterogeneous stress is white-noise random in the left column, whereas the smoother preferred model is shown in the right column.

Notice that there is a horizontal dotted line at a shear stress of 200 MPa. This signifies the yield stress at about 8.0 km (assumes Byerly-like friction). Grid points at which the shear stress already exceeds the yield stress are removed from the model. After that, time increases and the spatially uniform tectonic stress, $\dot{\boldsymbol{\sigma}}'^T (t - t_0)$, steadily increases at a rate of 10 kPa/yr. This stress rate is consistent with the strain rate inferred for the high-strain rate regions of Figure 8-68.

The amplitude of stress grows with time at most of the grid points in Fig.8-71. The **rate at which this stress amplitude grows is determined by the relative orientation of the tectonic stress and the initial stress**. If these two stresses

are aligned, then the growth rate of the amplitude is large. If, on the other hand the initial stress is opposite to the direction of the tectonic stress rate, then that point evolves ever further from the yield stress. Only grid points at which the stress amplitude exceeds the yield stress experience hypocenters. Once a grid point is declared to be a hypocenter, its location, origin time, and focal mechanism (orientation of maximum shear stress) are included in a synthetic catalog. Furthermore, once an event is generated, the location of that grid point is removed from the model. There is no concept of magnitude in this simulation; all events are simple grid points and there is no stress transfer to adjacent grid points when an event occurs. This means that there are no aftershocks in this model.
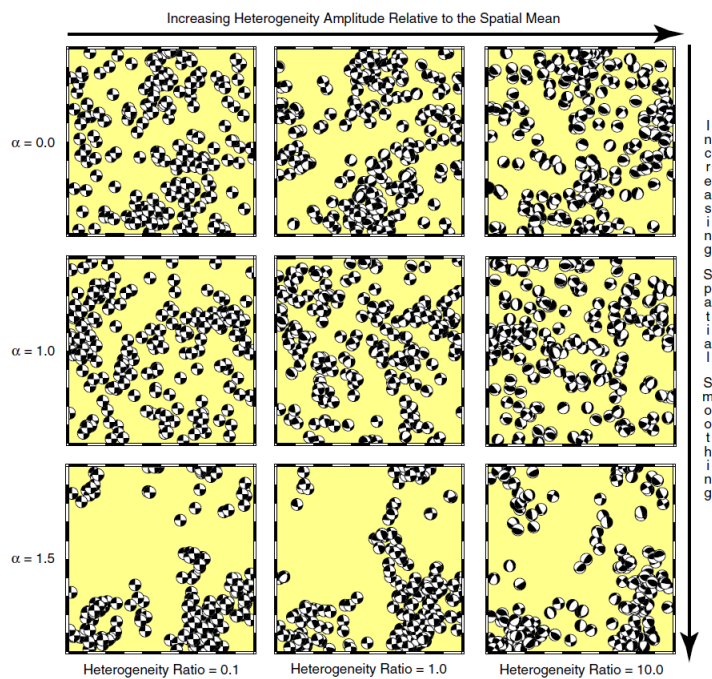
Figure 8-73. Simulated seismicity maps based on the fractal stress model of Smith and Heaton. Models increase in overall heterogeneity from left to right. Models increase in spatial roughness from bottom to top. Notice that spatial clustering increases as the models become smoother; the top row is random white noise plus a spatially uniform stress.

The most important concept in this model is that locations of hypocenters are **spatially biased; the seismicity is highest in places where the stochastic heterogeneous stress is aligned with the tectonic stress rate**. That is, one cannot use focal mechanism catalogs to infer the average stress orientation. Since focal mechanisms that are aligned with the stress rate tensor are favored, the inversion procedure tens to derive the orientation of the stress rate tensor. More importantly, **the strong spatial heterogeneity in the stress is invisible to focal mechanism catalogs** (except for aftershocks). Figure 8-73 shows examples of focal mechanism catalogs that are generated by the procedure just described. Maps of the seismicity produced by different roughnesses, α, and heterogeneity

ratios, HR, are shown. Notice that as α increases (i.e., the stochastic variation becomes smoother) the seismicity becomes more spatially clustered.

Figure 8-74 shows a comparison of the spatial coherence of focal mechanisms as measured by Hardebeck (2006) and a prediction of this value using the Smith and Heaton stochastic model. The horizontal axis is the distance between pairs of events, and the vertical axis is the average angular difference between the focal mechanisms of the pair. $\alpha = 0.75$ and $HR = 2.38$ produced simulations that were most similar to Hardebeck's data analysis. The Smith and Heaton study is difficult to describe in the simplified context of class notes. This is an important topic and I suggest that you attempt to read the entire manuscript (including Hardebeck's published comment and our reply).
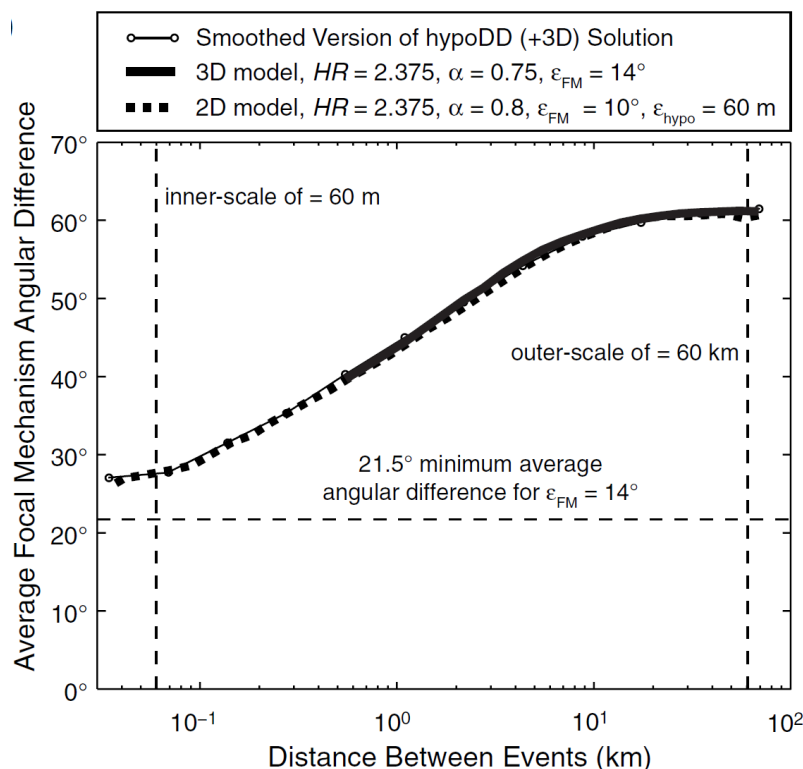
Figure 8-74. Comparison of the Smith and Heaton model for focal mechanisms in a stochastic stress field compared with the observation of Jeanne Hardebeck of a catalog of focal mechanisms in southern California (J. Hardebeck, 2006, Homogeneity of small-scale earthquake faulting, stress and fault strength, *Bulletin of the Seismological Society of America*, *96*, 1675-1688). The best fit to Hardebeck's analysis is with $\alpha = 0.8$ and $HR = 2.4$.

In order to get a better intuitive understanding of our best fit stochastic model, we generated an example of one component of the shear stress on a 100-km grid and at a spacing of 1 cm. Since there are too many grid points to observe on a single plot, we plotted a line through the grid at a variety of scales in Figure 8-75. In particular, panel (a) shows 10,000 points on a line through the entire 100-km volume, panel (b) shows the same stress on a 10-km length that run from 15 km to 25 km, which is a region of higher stochastic stress. Panel (c) focuses in on

17.5 km to 18.5 km, which is the largest stress patch in the previous panel. Finally, panel (d) focuses in on a 100-m patch of the highest stress in panel (c). The high shear stress at 18.01 km will cause an earthquake nucleation at that point at some time in the future.

Average shear stress at Different Lengths

We can use Figure 8-75 to investigate the stress-based strength of the crust based on this model (see Figure 8-52 and equation 8.122). In particular, the spatially averaged shear stress is shown for each panel in Fig. 8-75. This average shear stress is about 190 MPa at 1 m, 174 MPa at 100m, 155 MPa at 1 km, 104 MPa at 10 km, and 60 MPa at 100 km. It is clear that, for this stress distribution, the average stress near potential nucleation points decreases with the length over which the stress is averaged.
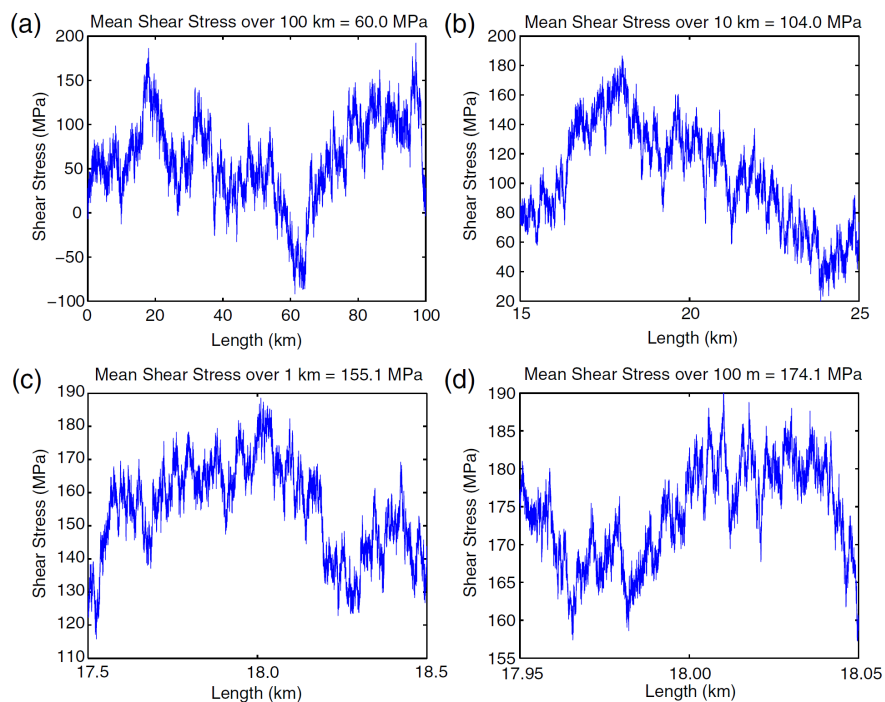
Figure 8-75 shows how the stress near a nucleation point (200 MPa yield) looks depending on the length scale. Clearly, the average stress in the longest length scale (panel a) in the 3-d fractal grid is lower than the average stress viewed at the shortest length scale (panel d).

Deborah Smith and I explored how this length scale dependence of average stress varies with the power-law smoothing parameter, α, using our stochastic stress model. The results of these numerical simulations are displayed in Figure 8-76. In this case, we assumed that

$$\overline{\sigma_L} \approx \left(\frac{1}{1+L} \right)^{\gamma}$$

8.129

where L is the scale length of the averaging and γ is a number determined from the stress generated in the model. Assuming that $\alpha \approx 0.8$ then indicates that the **average shear stress decreases as about the inverse fourth root of averaging length**. That is, changing from a length scale of 1 meter to a length scale of 10 km would decrease the average stress by a factor of ten.
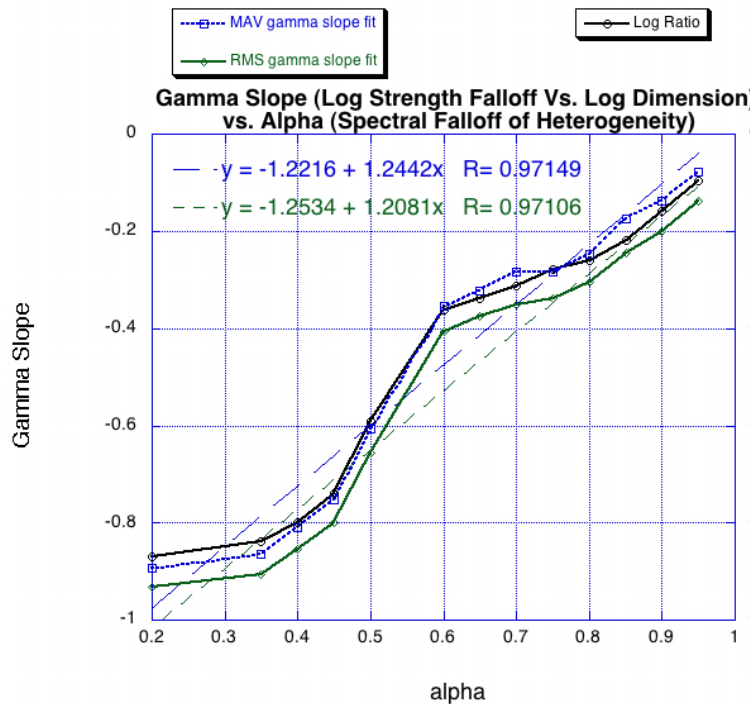
Figure 8-76. Based on simulations like that shown in Fig. 8-75, the dependence on stress-based strength on the length scale $\Gamma_\sigma \sim L^\gamma$ is shown for different values of fractal roughness.

There are numerous bold assumptions underlying this calculation and, unfortunately, it's almost impossible to directly test this model since the measurement of true stress is exceedingly difficult (i.e., expensive). The Cajon Pass Borehole experiment is one example of such a measurement. This was "the" major earthquake experiment of the late 1980's. It consisted of a 3.5 km deep borehole that was drilled about 3 km north of the San Andreas fault in Cajon Pass. With the view that rocks would be too damaged to measure if they were in the fault zone, the scientists managing the project chose to drill adjacent to the fault. A 300-m section of the downhole log of the orientation of borehole breakouts is shown in Figure 8-77. Panel (a) shows the observations. Note that at any given depth, there are two data points that are 180° different (breakouts on opposite sides of the borehole). Panel (b) shows the orientation of the principal compression axis along a 300-m line intersecting our preferred southern California stochastic stress model. Although the model and the observation are different, they do have similar variations over this length scale.

One of the most embarrassing observations of this expensive experiment was the fact that the spatial average of the orientation of the principal compression axis for the Cajon Pass borehole indicated that the San Andreas has an average shear

stress that is left lateral (at least in the Cajon Pass). This is embarrassing since the San Andreas fault is obviously a right-lateral fault. Furthermore, the geodetically measure strain rates are also right-lateral (see Figure 8-68). Barton and Zoback interpreted this observation to mean that the principal compression axis is almost perpendicular to the fault, which implies a very small shear stress on the fault. In our stochastic stress model, most regions are right-lateral. However, there are also significant regions of left-lateral stress.
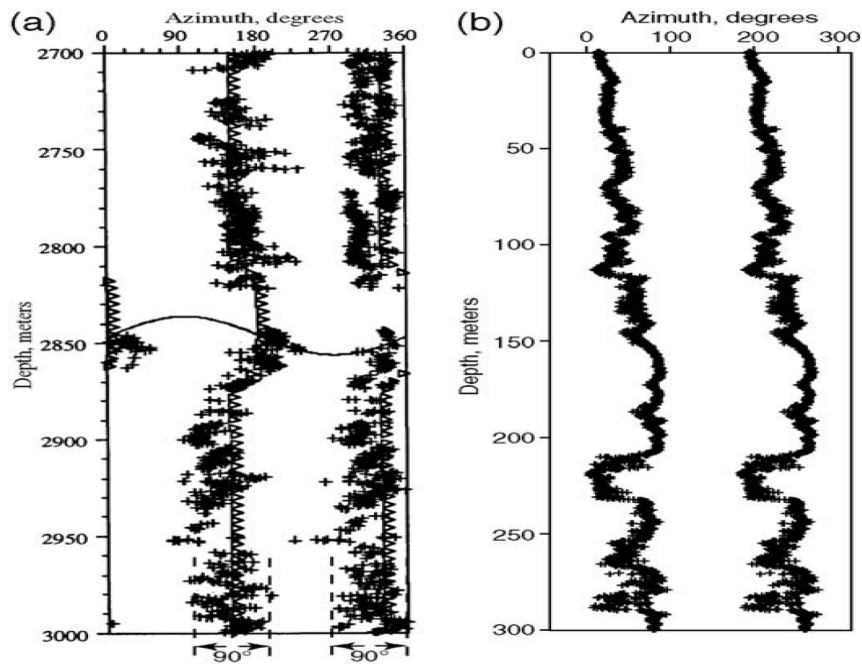
Figure 8-77. (a) orientation of borehole breakouts observed in a 300-m stretch of the Cajon Pass Borehole located 4 km from the San Andreas fault. Tensile stresses are maximum at opposite sides of a borehole and they are aligned with the maximum principal stress. (b) orientation of the maximum principal stress along a 300-m linew transecting the preferred fractal model of Smith and Heaton

The stochastic stress model of Smith and Heaton is radically different from other models of stress that are derived from assumptions that the stress variations are small compared to the average. Most researchers that I have encountered do not seriously entertain the possibility that shear stress could actually be opposite to the direction of tectonic strain accumulation. I must admit that the stochastic power-law model we proposed only makes sense if rupture dynamics are truly chaotic. While I personally believe that the evidence points us in that direction, the reader of these notes should always be cautious when discussing these issues (*caveat emptor*).

Failure Model that has low frictional heat, low fracture energy, and low radiated energy

Based on the insight obtained from the chaotic spring-block-slider model, the stochastic stress model of Smith and Heaton, and the gambler's ruin analysis of fault slip, I am now able to re-examine the stress paradox. That is, is there a model of dynamic rupture that has the following critical features? 1) high nucleation shear stress of about $0.6\sigma_n$, 2) small frictional heating (less than 2 MJ/m^2), 3) small fracture energy (less than 500 kJ/m^2, and 4) radiated far-field seismic energy of about $D \times 2.0 MJ/m^2$.

The key to satisfying these elements is to recognize that 1) friction depends very strongly on slip speed; low-speed sliding (mm/s) produces friction similar to the Byerly estimate, while high-speed sliding (> m/s) is almost frictionless. The transition between these two types of friction seems to occur when about 500 kJ/m^2 of energy is available on the sliding surface. 2) this low dynamic friction produces unsteady slip pulses that can propagate through regions of much lower average stress (even localized patches of negative stress). 3) the chaotic system evolves its own fractal-like prestress that is large enough to allow ruptures of any length (that is, the system is in a critical state).

The average stress drop of events depends on the distance that slip pulses propagate (heterogeneous slip is more likely to have smaller rupture lengths, and highly heterogeneous ruptures have average higher stress drops). Slip pulses propagate until they encounter low-stress regions. The Gutenberg-Richter relation is controlled by the spatial properties of the prestress. The prestress is the result of self-organization of a dynamic system that does not have inherent length scales.

One useful measure of this system is **seismic efficiency**, η_R, which is defined as the fraction of total energy that ends up as radiated seismic energy, or

$$\eta_R \equiv \frac{E_R}{\Delta W} = \frac{\Delta W - E_D}{\Delta W} = 1 - \frac{E_D}{\Delta W} \qquad 8.130$$

where E_R and E_D are radiated and dissipated energy, respectively. The observed low dissipation for large earthquakes implies that the seismic efficiency is high (approaching 1) for large earthquakes. That is, for large earthquakes,

$$\Delta W \Rightarrow E_{Radiation} = \sigma_{eff} P \approx 2 MPa \times P \qquad 8.131$$

We know that the change in potential energy is given by

$$\Delta W = \int_{-W/2}^{W/2} \int_{-L/2}^{L/2} \frac{1}{2} \left(\sigma_0(x,y) + \sigma_1(x,y) \right) D(x,y) \, dx \, dy$$

$$= \int_{-W/2}^{W/2} \int_{-L/2}^{L/2} \sigma_0(x,y) D(x,y) \, dx \, dy - \frac{1}{2} \int_{-W/2}^{W/2} \int_{-L/2}^{L/2} \Delta\sigma(x,y) D(x,y) \, dx \, dy$$

$$8.132$$

To get some idea of how things work, I will simplify things by assuming that the rupture is approximately a line source. That is, $D(x,y) \approx D(x)\delta(y)$. Note that if the rupture physics is isotropic, then x can be in any direction. I can now evaluate the change in potential energy as

$$\Delta W = \int_{-L/2}^{L/2} \sigma_0(x) D(x) \, dx - \frac{1}{2} \int_{-L/2}^{L/2} \Delta\sigma(x) D(x) \, dx \qquad 8.133$$

The first integral is a difficult problem since $D(x)$ is a highly nonlinear function of $\sigma_0(x)$. In contrast, the second integral is entirely determined by $D(x)$ since $\Delta\sigma(x)$ is a linear function of $D(x)$. In particular, it is convenient to write this in terms of a Green's function.

$$\Delta\sigma(x) = D(x) * G_{\Delta\sigma}(x) \qquad 8.134$$

Where $G_{\Delta\sigma}(x)$ is a Green's function for a spatial impulse of slip. Uenishi and Rice (2003, JGR, doi:10.1029/2001JB001681) say that

$$\Delta\sigma(x) = -\frac{\mu^*}{2\pi} \int_{-L/2}^{L/2} \frac{\frac{\partial D(\xi)}{\partial \xi}}{x - \xi} d\xi$$

$$= \frac{\mu^*}{2\pi} \int_{-\infty}^{\infty} \frac{\frac{\partial D(\xi)}{\partial \xi}}{\xi - x} d\xi \qquad 8.135$$

$$= \frac{\mu^*}{2\pi} \frac{\partial D(x)}{\partial x} * \frac{1}{x}$$

$$= -\frac{\mu^*}{2} F_{Hi}\left[\frac{\partial D(x)}{\partial x} \right]$$

Where $\mu^* = \mu$ for mode III and $\mu^* = \mu/(1-\nu)$ for modes I and II. F_{Hi} signifies a Hilbert transform. This can also be written

$$\Delta\sigma(x) = \frac{\mu^*}{2} \frac{\partial}{\partial x} F_{Hi}\left(D(x) \right) \qquad (8.136)$$

Therefore,

$$\Delta\sigma D = -D(x)\frac{\mu^*}{2\pi}F_{Hi}\big(D'(x)\big) \qquad 8.137$$

I can take the Fourier transform of 8.137 to obtain

$$FT(\Delta\sigma D) = \Delta\tilde{\sigma}(k) * \tilde{D}(k)$$

$$= -\frac{\mu^*}{2\pi}\tilde{D}(k) * F_{Hi}\big(ik\tilde{D}\big)$$

$$= -\frac{\mu^*}{2\pi}\tilde{D} * \big[-i^2 k\,\mathrm{sgn}(k)\tilde{D}\big] \qquad 8.138$$

$$= \frac{\mu^*}{2\pi}\big(k\,\mathrm{sgn}(k)\tilde{D}\big) * \tilde{D}$$

And then, using Parsevals' theorem

$$\int_{-L/2}^{L/2}\Delta\sigma D\,dx = \frac{\mu}{2\pi}\int_{-\infty}^{\infty}D(x)F_{Hi}\big(D'(x)\big)dx$$

$$= \frac{\mu}{2\pi}\int_{-\infty}^{\infty}\big[k\,\mathrm{sgn}(k)\tilde{D} * \tilde{D}\big]dk \qquad 8.139$$

$$\overset{?}{=}\frac{\mu}{2\pi}\int_{-\infty}^{\infty}\big[k\tilde{D}^2\big]dk$$

I now need to estimate $\tilde{D}^2(k)$. I begin with a statistical description of $D(x)$.

I will assume that at each point along the rupture, the slip is chosen from a random number generator with a pdf described by a mean-zero Normal distribution with a variance of σ^2. Meier, Ampuero and Heaton investigated the Potency rate functions that resulted from finite source inversions of large earthquakes. They found that

$$\dot{P}(t) \approx \left[\Pi\left(\frac{t}{2L/V_R}\right) * \Pi\left(\frac{t}{2L/V_R}\right)\right]\big(1-R(x)\big) \qquad 8.140$$

Where $R(x,\sigma=0.38)$ is a number chosen from a group of random numbers whose frequency (probability density function, or pdf) is described by a Gaussian function, $P = \dfrac{e^{-R^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$. That is, the random numbers are from a Normal distribution with a standard deviation $\sigma \approx 0.38$. 8.140 is an isosceles triangle (the expected value) multiplied by a random number that has a normal pdf with a mean of 1. Almost all of the random numbers are less than 1, so this multiplicative factor is almost entirely positive). It's important to realize that this functional description of spatially varying slip is based on fitting observed data (a rather complex observation); it's not clear that this functional form is derivable from physics.

Now recall (from Chapter 7) that, for a slip pulse propagating along a long narrow rupture (i.e., a line source), the Potency rate maps directly into the slip distribution. That is, $D(x) \approx \dot{P}(V_R t)$. From this, I infer that

$$D(x) \approx 6 \times 10^{-5} L \left[\Pi\left(\frac{x}{2L}\right) * \Pi\left(\frac{x}{2L}\right) \right] (1 - R(x)) \qquad 8.141$$

The scaling factor, $6 \times 10^{-5} L$, is from the observed ratios of average slip to rupture length (see 8.31). I can now attempt to calculate the Fourier transform of 8.141, which gives

$$8.142$$

Unfortunately, I am unable to take the Fourier Transform of 1. I note that I can write that

$$D^2(x) \approx E\left(R^2(x)\right) \left\{ 6 \times 10^{-5} L \left[\Pi\left(\frac{x}{2L}\right) * \Pi\left(\frac{x}{2L}\right) \right] \right\}^2 \qquad 8.143$$

Now the power of the random variable term is just the energy density of the spatially filtered random numbers. The final density is modulated in space by the triangle function. I think I can do my integral in the space domain by integrating the modulated power as a function of x.

$$\Delta W \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\tilde{\sigma}_0(k_x, k_y) - \frac{1}{2} \Delta \tilde{\sigma}(k_x, k_y) \right) * \tilde{D}^*(k_x, k_y) * \Pi(Lk_x) * \Pi(Wk_y) \, dk_x dk_y$$

$$8.144$$

$$\Delta W = \int_{-\infty}^{\infty} \left(\tilde{\sigma}_0(k_x) - \frac{1}{2} \tilde{D}(k_x) \tilde{G}_{\Delta\sigma}(k_x) \right) * \tilde{D}^*(k_x) * \Pi(Lk_x) \, dk_x \quad 8.145$$

Now the stress drop is a linear function of the slip distribution. That is, it can be calculated using a linear Green's function, $G_{\Delta\sigma}(x)$ or

$$\Delta\sigma(x) = D(x) * G_{\Delta\sigma}(x) \qquad 8.146$$

This convolution can also be defined in the frequency domain as

$$\Delta\tilde{\sigma}(k_x) = \tilde{D}(k_x) \tilde{G}_{\Delta\sigma}(k_x) \qquad 8.147$$

Unfortunately, I do not know what $G_{\Delta\sigma}(x)$ is, but for now,

The convolution with a rectangle is identical to performing a running-mean lo-pass filter (width $1/L$) on the integrand. Unfortunately, 8.145 is a rather cumbersome expression.

$$\Delta W = \int_{-\infty}^{\infty} \left[\tilde{\sigma}_0(k_x) - \frac{1}{2} \tilde{D}(k_x) \tilde{G}_{\Delta\sigma}(k_x) \right] * \tilde{D}^*(k_x) * \mathrm{sinc}(Lk_x) dk_x$$

$$\approx \int_{-\infty}^{\infty} \left[\tilde{\sigma}_0(k_x) * \mathrm{sinc}(Lk_x) * \tilde{D}^*(k_x) - \frac{1}{2} \tilde{D}(k_x) \tilde{G}_{\Delta\sigma}(k_x) * \tilde{D}^*(k_x) * \mathrm{sinc}(Lk_x) \right] dk_x$$

8.148

I can now take the Fourier transform of **Error! Reference source not found.**.

$$\tilde{G}_{\Delta\sigma}(k_x) = i\mu k_x \left(1 + \sqrt{k_x}\right) \hspace{2cm} 8.149$$

Where I used the fact that $FT\left(\sqrt{x}\right) = \sqrt{k_x}$. I am tired of explicitly writing everything as a function of k_x. I can now rewrite 8.148 as

$$\Delta W \approx \int_{-\infty}^{\infty} \left[\tilde{\sigma}_0 * \Pi(Lk_x) * \tilde{D}^* - \frac{1}{2} \mu i k_x \left(1 + \sqrt{k_x}\right) \tilde{D} * \tilde{D}^* * \Pi(Lk_x) \right] dk_x$$

$$= \int_{-\infty}^{\infty} \left[\tilde{\sigma}_0 * \Pi(Lk_x) * \tilde{D}^* - \frac{1}{2} \mu i k_x \left(1 + \sqrt{k_x}\right) \tilde{D}^2 * \Pi(Lk_x) \right] dk_x$$

8.150

At this point, I need to emphasize that we have little hope of obtaining accurate measurements of the prestress or the slip distribution, especially since we strongly expect that these are complicated functions of space. To deal with this, I will assume that these parameters can by described by their probability density functions. For the following section, you should understand that when I write prestress and slip, I actually mean that there are pdf's that describes the parameters for each value of x. I will then describe the distribution with an expected value and some pdf about the expected value. Now I will make the critical assumption that prestress and slip are random stationary functions of space. That is, I assume that statistical properties are invariant with respect to position. Furthermore, when I say the "change in potential energy," I really mean the *expected value of the change in potential energy*.

I need to be careful with 8.150, since it has terms that are the product of stochastic variables, that is, $\tilde{\sigma}_0(k_x) * \tilde{D}^*(k_x)$ and also $\tilde{D}^*(k_x)\tilde{D}^*(k_x)$. These products are also stochastic variables, with the following properties

$$E\left[\tilde{\sigma}_0(k_x)\tilde{D}(k_x)\right] = E\left[\tilde{\sigma}_0(k_x)\right]E\left[\tilde{D}(k_x)\right] + \mathrm{cov}\left[\tilde{\sigma}_0(k_x), \tilde{D}(k_x)\right]$$

8.151

And

$$E\left[\tilde{D}^*(k_x)\tilde{D}(k_x)\right] = E\left[\tilde{D}^*(k_x)\right]E\left[\tilde{D}(k_x)\right] + \mathrm{cov}\left[\tilde{D}^*(k_x), \tilde{D}(k_x)\right]$$

$$= E^2\left[\tilde{D}^*(k_x)\right] + \mathrm{var}\left[\tilde{D}(k_x)\right]$$

8.152

Where cov is the covariance function, which is defined as

$$\text{cov}\left[\tilde{\sigma}_0(k_x), \tilde{D}(k_x)\right] \equiv E\left[\left(\tilde{\sigma}_0(k_x) - E\left[\tilde{\sigma}_0(k_x)\right]\right)\left(\tilde{D}(k_x) - E\left[\tilde{D}(k_x)\right]\right)\right]$$

$$8.153$$

Covariance is a measure of the correlation between random variables; when the covariance is positive, then two variables tend to vary together (positive variations in one variable tend to occur when there are positive variations in the other variable. When the variables are statistically independent, then their covariance is zero. I am getting tired of explicitly writing that my random variables are a function of k_x, so I will drop that formality. I can obtain the expected change in strain energy by substituting 8.151 and 8.152 into 8.150.

$$E[\Delta W] \approx \int_{1/L}^{\infty} \left[2E[\tilde{\sigma}_0]E[\tilde{D}^*] + 2\,\text{cov}\left(\tilde{\sigma}_0, \tilde{D}^*\right) - \left\{E[\tilde{D}^2] + \text{var}\left(\tilde{D}^*\right)\right\} \mu i k_x \left(1 + \sqrt{k_x}\right)\right] dk_x$$

8.154

At this point, I will assume that my random variables, σ_0 and D, are a *zero-mean, Gaussian, stationary process*. That is, at each value of x, I assign a value by picking a random number, where the probability of the number is given by the Gaussian function. Basically, I start with Gaussian white noise and I then apply a spatial filter power-law that is a power law in wavenumber space (I spatially smooth the random white noise). I now note that the Fourier transform of Gaussian white noise is Gausian white noise. Furthermore, it is easy to show that

$$E\left(\tilde{D}^2\right) = \text{var}\left(\tilde{D}\right) + E^2\left(\tilde{D}\right) \qquad\qquad 8.155$$

Since I assumed that $D(x)$ and $\sigma_0(x)$ are mean zero, then it follows that $\tilde{D}(k)$ and $\tilde{\sigma}_0(k_x)$ are also mean zero. Therefore 8.154 becomes

$$E[\Delta W] \approx 2\int_{1/L}^{\infty} \left[\text{cov}\left(\tilde{\sigma}_0, \tilde{D}^*\right) - \text{var}\left(\tilde{D}^*\right) \mu i k_x \left(1 + \sqrt{k_x}\right)\right] dk_x \quad 8.156$$

Now the covariance term can be written as

$$\text{cov}\left(\tilde{\sigma}_0, \tilde{D}\right) = \text{corr}\left(\tilde{\sigma}_0, \tilde{D}\right)\sqrt{\text{var}\left(\tilde{\sigma}_0\right)\text{var}\left(\tilde{D}\right)} \qquad\qquad 8.157$$

Where corr is the correlation coefficient between prestess and the slip. A correlation of +1 means that variables are perfectly correlated, -1 if they are oppositely correlated, and 0 if they are statistically independent. So 8.156 can be written

$$E[\Delta W] \approx 2\int_{1/L}^{\infty} \left[\text{corr}\left(\tilde{\sigma}_0, \tilde{D}\right)\sqrt{\text{var}\left(\tilde{\sigma}_0\right)\text{var}\left(\tilde{D}\right)} - \text{var}\left(\tilde{D}^*\right) \mu i k_x \left(1 + \sqrt{k_x}\right)\right] dk_x$$

$$8.158$$

At this point, I will assume that my variables are wave-number filtered versions of Gaussian white noise with a variance of 1. That is, I assume that

$$\mathrm{var}\left(\tilde{\sigma}_0\right) \approx \left(\frac{1}{\left(1+k_x\right)^{\beta}}\right)^2 \qquad\qquad 8.159$$

And

$$\mathrm{var}\left(\tilde{D}\right) \approx \left(\frac{1}{\left(1+k_x\right)^{\alpha}}\right)^2 \qquad\qquad 8.160$$

Substituting these into 8.158

$$E\left[\Delta W\right] \approx 2\int_{1/L}^{\infty}\left[\mathrm{corr}\left(\tilde{\sigma}_0,\tilde{D}\right)\frac{1}{\left(1+k_x\right)^{\beta}} - \mu k_x\left(1+\sqrt{k_x}\right)\right]\frac{1}{\left(1+k_x\right)^{\alpha}}dk_x$$

$$= 2\mathrm{corr}\left(\tilde{\sigma}_0,\tilde{D}\right)\int_{1/L}^{\infty}\frac{1}{\left(1+k_x\right)^{\beta}}dk_x - \mu\int_{1/L}^{\infty}\frac{k_x + k_x^{3/2}}{\left(1+k_x\right)^{\alpha}}dk_x$$

$$= 2\mathrm{corr}\left(\tilde{\sigma}_0,\tilde{D}\right)\left|\frac{\left(1+k_x\right)^{\beta+1}}{\beta+1}\right|_{1/L}^{\infty} - \mu\left|\frac{\left(1+k_x\right)^{1-\alpha}\left[\left(\alpha-1\right)k_x+1\right]}{\left(\alpha-2\right)\left(\alpha-1\right)} + \frac{k_x^{5/2-\alpha}}{5/2-\alpha}\right|_{1/L}^{\infty}$$

$$8.161$$

can simplify it by assuming that the prestress, slip, and stress drop are isotropic on the fault plane. That is, I will assume that

$$D\left(x,y\right) \approx D_r\left(r\right) \qquad\qquad 8.162$$

Where $r = \sqrt{x^2 + y^2}$. At this point, I need to modify my 2-d Fourier transform such that it only depends on r. In particular,

$$\tilde{D}\left(k_x,k_y\right) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}D\left(x,y\right)e^{-i2\pi\left(xk_x+yk_y\right)}dxdy$$

$$= \int_{0}^{\infty}D_r\left(r\right)\left[\int_{0}^{2\pi}e^{-i2\pi\left(rk_r\cos\theta\right)}d\theta\right]rdr \qquad\qquad 8.163$$

$$= 2\pi\int_{0}^{\infty}D_r\left(r\right)J_0\left(2\pi rk_r\right)rdr$$

$$\equiv \hat{D}_r\left(k_r\right)$$

That is, I will greatly simplify the problem by assuming that the Fourier spectrum of slip is approximately given by

$$\tilde{D}\left(k_x, k_y\right) \approx \hat{D}_r\left(k_r\right) \qquad 8.164$$

Where the **hat connotes Hankel transform of zero order**, J_0 is a Bessel function of the first kind, and $k_r \equiv \sqrt{k_x^2 + k_y^2}$. A Hankel transform is very similar to our familiar Fourier transform, except that the kernel function is a Bessel function instead of a sinusoid. Similar to a sinusoid, $J_0(r)$ oscillates with increasing r. However a Bessel function's amplitude decays with distance and the asymptotic expansion is

$$J_0\left(r\right) \approx \sqrt{\frac{2}{\pi r}} \cos\left(r - \frac{\pi}{4}\right) \qquad 8.165$$

The Hankel transform has the following properties

$$\hat{D}\left(k_r\right) \equiv 2\pi \int_0^\infty D_r\left(r\right) J_0\left(2\pi r k_r\right) dr \qquad 8.166$$

$$D_r\left(r\right) = 2\pi \int_0^\infty \hat{D}_r\left(k_r\right) J_0\left(2\pi r k_r\right) k_r dk_r \qquad 8.167$$

This is an unrealistic assumption that allows me to simplify the math. In particular, 8.145 becomes

$$\Delta W = \int_{-\infty}^\infty \int_0^{2\pi} \left(\tilde{\sigma}_0\left(k_r\right) - \frac{1}{2}\tilde{D}\left(k_r\right)\tilde{G}\left(k_r\right)\right)\tilde{D}\left(k_r\right)\mathrm{sinc}\left(Rk_r\right)\left(k_r^{-1}d\theta\right)dk_r$$

$$= 2\pi \int_{-\infty}^\infty \left(\tilde{\sigma}_0\left(k_r\right) - \frac{1}{2}\tilde{D}\left(k_r\right)\tilde{G}\left(k_r\right)\right)k_r^{-1}\tilde{D}\left(k_r\right)\mathrm{sinc}\left(Rk_r\right)dk_r$$

$$8.168$$

Now I will note that a sinc function with scale width of R can be crudely approximated by a Rectangle function with a scale width of R. 8.168 can be approximated as

$$\Delta W = 2\pi \int_{-\infty}^\infty \left(\tilde{\sigma}_0\left(k_r\right) - \frac{1}{2}\tilde{D}\left(k_r\right)\tilde{G}\left(k_r\right)\right)k_r^{-1}\tilde{D}\left(k_r\right)\Pi\left(Rk_r\right)dk_r$$

$$= 2\pi \int_0^{1/R} \left(\tilde{\sigma}_0\left(k_r\right) - \frac{1}{2}\tilde{D}\left(k_r\right)\tilde{G}\left(k_r\right)\right)k_r^{-1}\tilde{D}\left(k_r\right)dk_r$$

$$\qquad 8.169$$

$$= 2\pi \int_0^{1/R} \frac{\tilde{\sigma}_0\left(k_r\right)\tilde{D}\left(k_r\right)}{k_r}dk_r - \pi \int_0^{1/R} \frac{\tilde{D}^2\left(k_r\right)\tilde{G}\left(k_r\right)}{k_r}dk_r$$

$$= 2\pi \int_0^{1/R} \frac{\tilde{\sigma}_0\left(k_r\right)\tilde{D}\left(k_r\right)}{k_r}dk_r - \pi \int_0^{1/R} \frac{\tilde{D}\left(k_r\right)\Delta\tilde{\sigma}\left(k_r\right)}{k_r}dk_r$$

At this point, I will assume that the prestress and the slip are both random variables with pdf's given by

$$\tilde{\sigma}_0\left(k_r\right) \approx C_{\sigma_0} k^{-\alpha} \qquad 8.170$$

$$\tilde{D}(k_r) \approx C_D k^{-\beta} \qquad\qquad 8.171$$

I can now approximate the change in potential energy 8.132 as

$$\Delta W = 2\pi \int_0^{1/R} E\left[\frac{\tilde{\sigma}_0(k_r)\tilde{D}(k_r)}{k_r}\right] dk_r - \pi \int_0^{1/R} E\left[\frac{\tilde{D}(k_r)\Delta\tilde{\sigma}(k_r)}{k_r}\right] dk_r$$

$$= 2\pi \int_0^{1/R} \left\{ E\left[\tilde{\sigma}_0(k_r)\right] E\left[\tilde{D}(k_r)\right] + \mathrm{cov}\left[\tilde{\sigma}_0(k_r),\tilde{D}(k_r)\right] \right\} k_r^{-1} dk_r - \pi \int_0^{1/R} E\left[\tilde{D}(k_r)\Delta\tilde{\sigma}(k_r)\right] k_r^{-1} dk_r$$

$$8.172$$

$$\Delta W = \iint_S \left(\sigma_0(x,y) - \frac{1}{2}\mu f(y)\frac{\partial}{\partial x}D(x)\right) D(x)f(y)\,dxdy$$

$$8.173$$

$$= \iint_S \left(\sigma_0(x,y)f(y)D(x) - \frac{1}{4}\mu f^2(y)\frac{\partial}{\partial x}D^2(x)\right) dxdy$$

For the moment, I will greatly simplify things by assuming that $f(y) \approx \Pi(y)$. I can now rewrite 8.173 as

$$\Delta W \approx W \int_0^L \left(\sigma_0(x)D(x) - \frac{1}{4}\mu\frac{\partial}{\partial x}D^2(x)\right) dx \qquad 8.174$$

The observation that the ratio of $\bar{D}/L \approx 10^{-5}$ can be produced by a random-phase slip that has an amplitude spectrum of $\tilde{D}(k) = k^{-\alpha}$, where $\alpha \approx 1.2$ (see Zheng-Liu and others). From 8.130,

$$\Delta W = \frac{E_R}{\eta} \qquad\qquad 8.175$$

So 8.175 is rewritten as

$$E_R \approx \eta W \int_0^L \left(\sigma_0(x)D(x) - \frac{1}{4}\mu\frac{\partial}{\partial x}D^2(x)\right) dx \qquad 8.176$$

Now radiated energy is estimated to be $E_R \approx \sigma_{eff}P$ (see 8.7), so 8.176 becomes

$$E_R \approx \sigma_{eff}P \approx \eta W \int_0^L \left(\sigma_0(x)D(x) - \frac{1}{4}\mu\frac{\partial}{\partial x}D^2(x)\right) dx \qquad 8.177$$

Or

$$\sigma_{eff} W \int_0^L D(x)\, dx \approx \eta W \int_0^L \left(\sigma_0(x) D(x) - \frac{1}{4} \mu \frac{\partial}{\partial x} D^2(x) \right) dx \qquad 8.178$$

$$\sigma_{eff} \int_0^L \left[D(x) + \frac{1}{4} \mu \frac{\partial}{\partial x} D^2(x) \right] dx \approx \eta \int_0^L \left(\sigma_0(x) D(x) \right) dx$$

$$\sigma_{eff} \int_0^L \left[D(x) + \frac{1}{2} \mu D(x) \frac{\partial}{\partial x} D(x) \right] dx \approx \eta \int_0^L \left(\sigma_0(x) D(x) \right) dx$$

$$\sigma_{eff} \int_0^L \left[1 + \frac{1}{2} \mu \frac{\partial}{\partial x} D(x) \right] D(x)\, dx \approx \eta \int_0^L \left(\sigma_0(x) D(x) \right) dx \qquad 8.179$$

$$\Rightarrow \sigma_{eff} D(x) \left[1 + \frac{1}{2} \mu \frac{\partial}{\partial x} D(x) \right] = \eta \sigma_0(x) D(x)$$

$$\sigma_{eff} \left[1 + \frac{1}{2} \mu \frac{\partial}{\partial x} D(x) \right] = \eta \sigma_0(x)$$

$$E_D = \iint_S \sigma_F\left(x, y; D, \dot{D}\right) D(x, y)\, dx dy \qquad 8.180$$

IF we assume that the prestress and final stress are uniform in x and y, then 8.132 becomes

$$\Delta W_{uniform} = \frac{1}{2}\left(\sigma_0 + \sigma_1\right) \bar{D} S \qquad 8.181$$

Of course the prestress and final stress can only be uniform for one unique distribution of slip which depends on the geometry of the rupture surface. If we make the very restrictive assumption that the friction is uniform and equal to the final stress, then

$$E_D = \sigma_1 \bar{D} S + E_G \qquad 8.182$$

Where E_G is fracture energy. If we assume that the radiated energy is the change in potential energy minus the dissipation energy, then in the case of uniform stress and friction,

$$E_R = \frac{1}{2}\left(\sigma_0 + \sigma_1\right) \bar{D} S - \sigma_1 \bar{D} S - E_G = \frac{1}{2}\left(\sigma_0 - \sigma_1\right) \bar{D} S - E_G$$

$$= \frac{1}{2}\left(\Delta\sigma\right) \bar{D} S - E_G$$

$$\equiv \Delta W_0 - E_G$$

$$8.183$$

Where E_G is the fracture energy, which is meant to signify the transition from static to dynamic friction (sometimes called the breakdown energy). In the case

of slip weakening friction, $E_G = \frac{1}{2}\left(\sigma_{yield} - \sigma_1\right)D_0 S$. Assuming slip weakening friction, the potential energy change available for seismic radiation is $\Delta W_0 \equiv \frac{1}{2}\Delta\sigma\bar{D}S - E_G$. Unfortunately, it is not possible for seismologists to determine absolute stress, we can only infer stress changes. In response, many seismologists have substituted ΔW_0 for ΔW in their definition of seismic efficiency. I will call this "crack seismic efficiency" (η_{crack}). In this case

$$\eta_{crack}^{observed} \equiv \frac{E_R^{observed}}{\Delta W_0} = \frac{2E_R^{observed}}{\Delta\sigma\bar{D}S} \qquad 8.184$$

We also know that

$$\begin{aligned}\eta_{crack} &= 1 - \frac{E_G}{\Delta W_0} \\ &= 1 - \frac{\left(\sigma_{yield} - \sigma_1\right)D_0}{\Delta\sigma\bar{D}}\end{aligned} \qquad 8.185$$

Or

$$D_0 = \frac{\left(1 - \eta_{crack}\right)\Delta\sigma\bar{D}}{\sigma_{yield} - \sigma_1} \qquad 8.186$$

In principle, seismologically observed quantities can be used to determine everything in 8.184. Then assuming that the Byerly friction is the yield stress, then the slip weakening distance can be determined, that is provided that σ_1 is known (which it is not). However, in this case, $E_G^{crack} = \left(1 - \eta_{crack}\right)\Delta\sigma\bar{D}S$.

Equations 8.181 through 8.186 all assume that the frictional sliding stress is the final stress after all rupture has ceased. The resulting definition of the change in potential energy is **very, very** restrictive. It assumes constant friction and it thus **excludes slip pulses**. Furthermore, all of these definitions assume that the prestress and stress drop are spatially uniform. This is clearly not the case. In particular, the slip and the stress drop are both complex functions of space that are correlated. Multiplying their average values is clearly inappropriate. The best description of the radiated energy is to subtract 8.180 from 8.132, or

$$E_R = \iint_S \left[\sigma_0(x,y) - \frac{1}{2}\Delta\sigma(x,y) - \sigma_F(D,\dot{D})\right]D(x,y)\,dxdy$$

8.187

As you are aware, I believe that the prestress and stress drop is very heterogeneous. Unfortunately, I don't know any simple average relations between radiated energy, slip and stress drop.

With regards to the fracture energy, I like the definition that $E_G \equiv \left(E_f - D\sigma_{\min} \right) dS$. This is rather arbitrary. However, the key issue about dynamics is whether or not the energy available to radiate increases with additional rupture. That is, if $\dfrac{dE_R}{dt} = 0$, then the event has ended. This can be accomplished solely with the distribution of prestress, which then determines the distributions of friction and slip. One does not need large fracture energies to stop an event. If the friction is strongly rate weakening, it produces slip pulses and the relationship between slip and dynamic friction and stress drop becomes very complex (see Elbanna and Heaton).

One potential research direction is to use seismological evidence to determine $E_R^{observed}$, $D(x,y)$, and $\Delta\sigma(x,y)$. We could then assume a friction law together with the assumption that slip is a Joffe slip pulse to determine $E_F(x,y) = F_{Joffe}\left(D(x,y) \right)$. This is similar to the approach taken in Elbanna and Heaton. In this case we can anticipate that $E_F \approx CD^\nu$, where C and ν are constants obtained from numerical simulations of pulses. Then 8.187 could be written

$$E_R^{observed} = \iint_S \sigma_0(x,y) D(x,y)\,dxdy - \iint_S \left\{ CD(x,y)^\nu + \frac{1}{2}\Delta\sigma(x,y)D(x,y) \right\} dxdy$$

$$= \iint_S \left\{ \left[\sigma_0(x,y) - \frac{1}{2}\Delta\sigma(x,y) \right] D(x,y) - CD(x,y)^\nu \right\} dxdy$$

$$8.188$$

Finally, we can constrain put a constraint on the prestress.

$$\iint_S \sigma_0(x,y) D(x,y)\,dxdy = E_R^{observed} + \iint_S \left\{ \frac{1}{2}\Delta\sigma(x,y) D(x,y) + CD(x,y)^\nu \right\} dxdy$$

$$8.189$$

The system could work consistently even with no friction or fracture energy as long as

$$E_R^{observed} = \iint_S \left\{ \left[\sigma_0(x,y) - \frac{1}{2}\Delta\sigma(x,y) \right] D(x,y) \right\} dxdy \qquad 8.190$$

As an example we use the relationship between radiated energy and seismic moment of Kanamori. We recall that, on average,

$$E_R \approx \frac{M_0}{2 \times 10^4} = \frac{\mu \overline{D} S}{2 \times 10^4} = \sigma_e \overline{D} S = \sigma_e P \qquad 8.191$$

Where effective energy $\sigma_e \approx 2 MPa$. Substituting this into 8.188, we obtain

$$\sigma_e S \bar{D} \approx \iint_S \left\{ \left[\sigma_0(x,y) - \frac{1}{2}\Delta\sigma(x,y) \right] D(x,y) - CD(x,y)^v \right\} dxdy$$

<div align="center">8.192</div>

If we assume that $\frac{\partial}{\partial D}\left[\sigma_0(x,y) - \frac{1}{2}\Delta\sigma(x,y) \right] = 0$, or that the final slip at a

point is independent from the difference between the initial stress and the stress drop at the point (implausible assumption), then 8.192 becomes,

$$\sigma_e \approx \overline{\left[\sigma_0(x,y) - \frac{1}{2}\Delta\sigma(x,y) \right]} - \frac{E_F}{P} \qquad\qquad 8.193$$

If $\overline{\Delta\sigma} \approx 3MPa$, then $\overline{\sigma_0(x,y)} \approx \sigma_e + 1.5MPa + \frac{E_F}{P} \approx 3.5MPa + \frac{E_F}{P}$. The fact

that there is minimal melting implies that $\frac{E_F}{P} < 2Mpa$. then $\overline{\sigma_0(x,y)} < 5.5MPa$.

This is a restatement of the stress paradox. Stresses inferred from seismology are far smaller than stresses measured in the laboratory, and they are small compared to the stresses required to resist gravity in mountain ranges. I am convinced that the key to resolving these problems is to establish the detailed connection between prestress and stress drop. In dynamic models that produce slip pulses (strong rate weakening friction), the amplitude of the slip pulse varies in a complex way that is mostly determined by the prestress and the hypocentral location (see the pulse energy equation by Elbanna and Heaton).

Strength-length scaling
b-value and prestress
b-value and brittleness
creeping fault
discrete vs continuum
far from equilibrium Reynolds number

A window covering the extent of one of the events
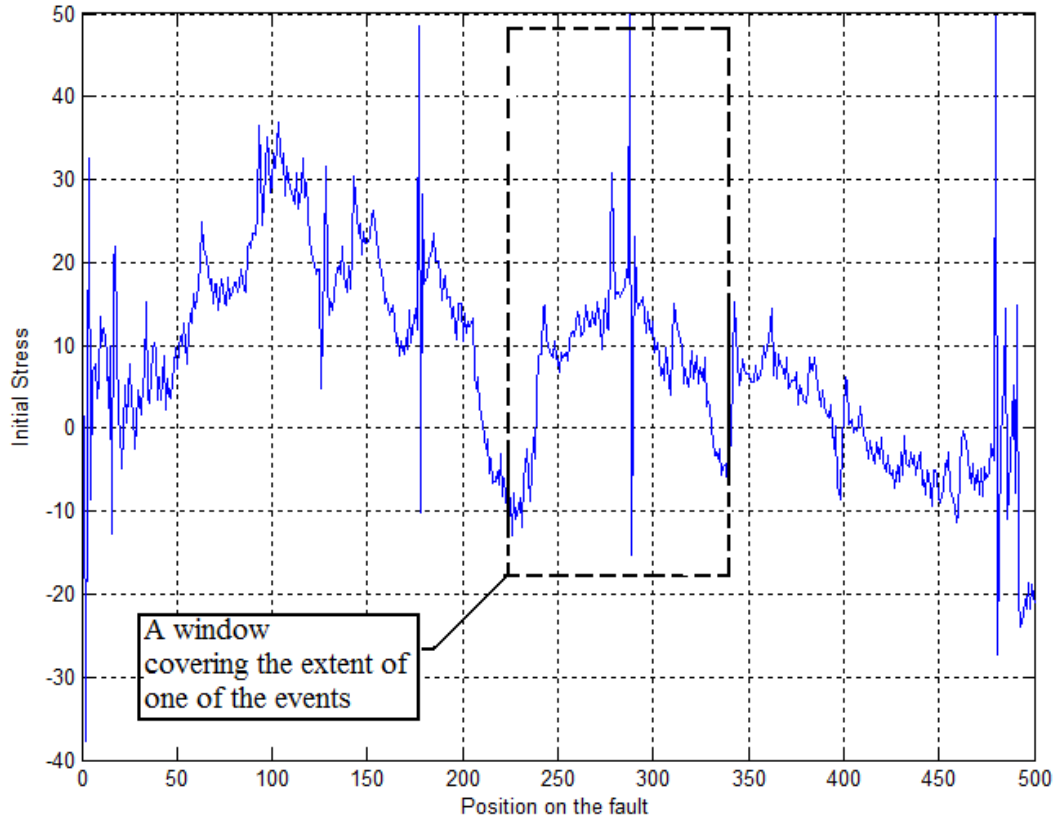
A Test Case

Prestress as a function of position

100-10,000 Times Faster!
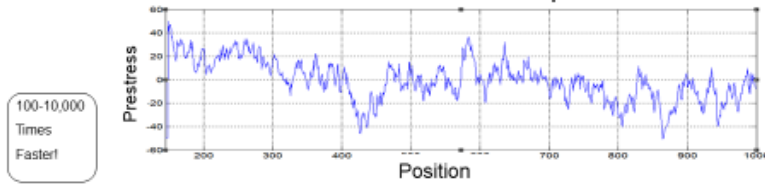
Slip as a function of position for both the spring block model and the proposed ODE
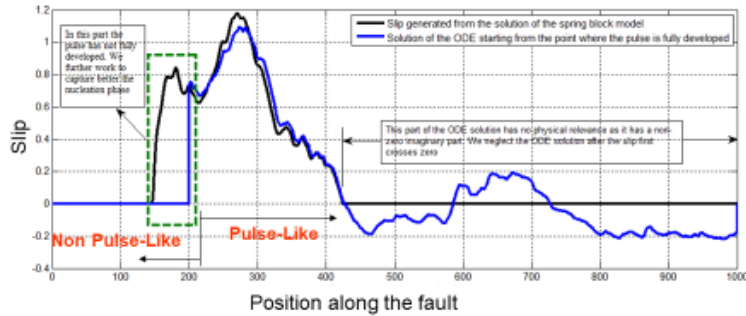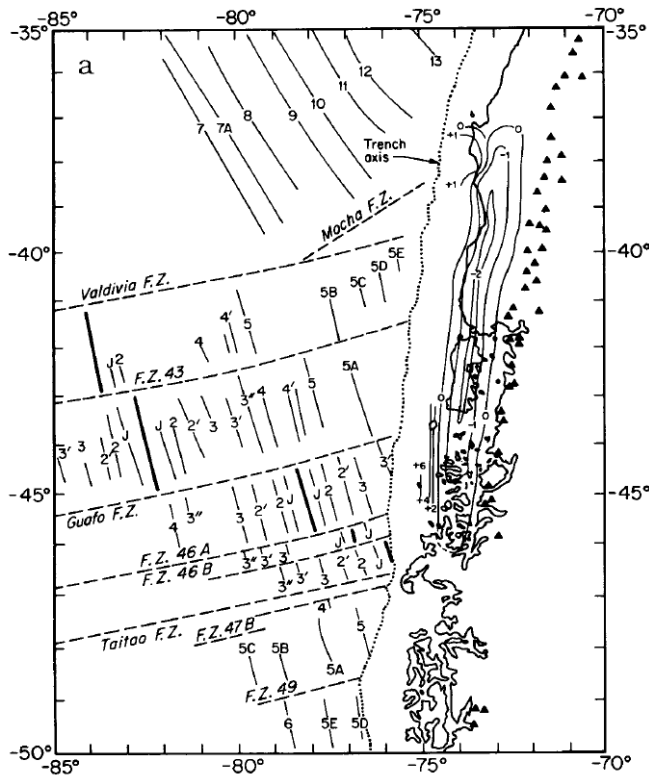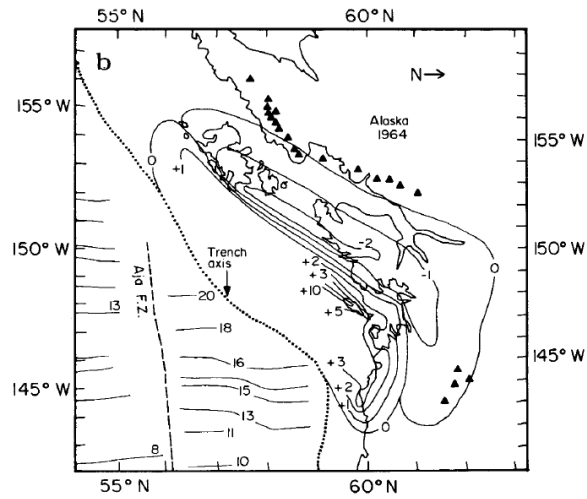
Derivation of Brune spectrum from the observation of near-source high-frequency magnitude saturation

We observe that very near-source accelerations (high-frequencies) appear to be incoherent noise with a peak acceleration that is independent of magnitude. This leads us to the following hypothesis.

Hypothesis: Radiated high-frequency energy $E_R^{\omega \gg \omega_c}$ scales with the rupture area S, independent of the average slip on that rupture surface, or

$$E_R^{\omega \gg \omega_c} \sim S \qquad\qquad 8.2$$

This hypothesis means that if we double the rupture area, then we double the radiated high-frequency energy.

What does this mean for the scaling of radiated seismic waves?

Let us **assume** that the **seismic wave is approximately a non-negative function of duration** T_c, whose **integrated area scales with seismic moment, and which has a power-law high-frequency spectral decay**. Then,

$$\tilde{U}_R \sim M_0 \text{sinc}^\alpha \left(\omega / \omega_c \right) \qquad\qquad 8.3$$

Where $\omega_c = 2\pi / T_c$. Equation **Error! Reference source not found.** has asymptotes

$$\tilde{U}_R \sim \begin{cases} M_0 & \omega \ll \omega_c \\ M_0 \left(\omega / \omega_c \right)^{-\alpha} & \omega \gg \omega_c \end{cases} \qquad\qquad 8.194$$

Now let us **assume that the duration of the signal is proportional to the dimension of the fault,** or that $\omega_c \sim \sqrt{S}$. In addition, **we will assume that** $M_0 \sim S\bar{D} \sim S^{3/2}$. Therefore, $\omega_c \sim M_0^{-1/3}$, and 8.194 becomes

$$\tilde{U}_R \sim \begin{cases} M_0 & \omega \ll \omega_c \\ M_0^{1-\alpha/3} \omega^{-\alpha} & \omega \gg \omega_c \end{cases} \qquad\qquad 8.195$$

Now the radiated energy spectrum scales as the square of the Fourier amplitude spectrum, or

$$E_R^{\omega \gg \omega_c} \sim \left(\tilde{U}_R^{\omega \gg \omega_c} \right)^2 \sim M_0^{2\left(1-\alpha/3\right)} \omega^{-2\alpha} \qquad\qquad 8.196$$

Now our original hypothesis was that

$$E_R^{\omega \gg \omega_c} \sim S \sim M_0^{2/3} \qquad\qquad 8.197$$

Relations 8.196 and 8.197 can only be simultaneously true if $\alpha = 2$. Therefore our hypothesis that the high-frequency very near-source ground motion is incoherent noise of constant amplitude that is independent of the size of the slip implies that

$$\tilde{U}_R \sim M_0 \text{sinc}^2 \left(\omega / \omega_c \right) \sim \begin{cases} M_0 & \omega \ll \omega_c \\ M_0^{1/3} \omega^{-2} & \omega \gg \omega_c \end{cases} \qquad\qquad 8.198$$

Relationship 8.198 is identical to the Brune spectrum (1970), but without the stress drop scaling. That is, the high-frequency radiation is independent of the stress drop. Therefore our hypothesis is identical to assuming Brune's spectral scaling for seismic moment (but not for stress drop).

We can also anticipate the following asymptotic behavior for any ground motion prediction equations. When the distance is large compared to the source dimension, and when the predominant periods of the ground motion are large compared to the source duration, we expect the ground motion amplitudes to scale with the seismic moment, or

$$\log U_{far\&lowfreq} \sim \log M_0 \sim \frac{3}{2}M \qquad \text{8.199}$$

For very near-source long-periods (e.g. displacement), we expect the peak amplitude to scale with the size of the slip on the nearby fault segment, or

$$\log U_{near\&lowfreq} \sim \log \bar{D} \sim \log M_0^{\frac{1}{3}} \sim \frac{1}{2}M \qquad \text{8.200}$$

Gutenberg-Richter Frequency Magnitude Relation

$$\log(N > M) = a - bm \qquad \text{8.201}$$

$$\frac{d(N > M)}{dM} = \frac{d}{dM}\left(10^{a-bM}\right) = -b\ln(10)10^{a-bM} \qquad \text{8.202}$$

$$N'(M - \frac{\Delta M}{2} < M < M + \frac{\Delta M}{2}) = -b\ln(10)10^{a-bM}\Delta M \qquad \text{8.203}$$

$$\log\left[N'(M - \frac{\Delta M}{2} < M < M + \frac{\Delta M}{2})\right] = \log\left[\ln(10)^{-b}\Delta M\right] + a - bM$$
$$\text{8.204}$$

$$N'(M - \frac{\Delta M}{2} < M < M + \frac{\Delta M}{2}) = 10^{\log\left[\ln(10)^{-b}\Delta M\right]+a-bM} = C''10^{-bM} \quad \text{8.205}$$

$$M = C - \frac{2}{3}\log M_0 = C' + \log\left(L^2 D\right)^{-\frac{2}{3}} \qquad \text{8.206}$$

If $L \sim D$, then

$$M = C'' + \log L^{-2} \qquad \text{8.207}$$

$$N'(M - \frac{\Delta M}{2} < M < M + \frac{\Delta M}{2}) = C''10^{-b\left(C''+\log L^2\right)} = \dot{C}L^{-2b} \qquad \text{8.208}$$

Now the total rupture area $A_{total}(M,\Delta M)$ for all earthquakes with $M - \frac{\Delta M}{2} < M < M + \frac{\Delta M}{2}$ is

$$A_{total}(M,\Delta M) = N'L^2 = \dot{C}L^{2(1-b)} \qquad \text{8.209}$$

That is, if the b-value is 1, then the total rupture area is the same for integrated area of each magnitude. That is the sum of the rupture area of all 2's is the same as the rupture area of all 3's, is the same as all 4's, etc. What this means is that

given a b-value of 1, and given that a point has just experienced slip, then it is equally likely that it could have come from any magnitude earthquake. Given that different magnitude earthquakes have different slips, any slip is as likely as any other. This is only true in a logarithmic sense. That is, a fault is equally likely to experience slip between D and C x D, regardless of the value of D and a constant, C (e.g. given, slips between 1 and 2 mm are just as common as slips between 4 and 8 m.

but with a mirror image fault
with plan view of displacement amplitude vs. distance from fault.
cross section of static motion of a dip-slip fault.
Static offsets in a half-space
near-fault particle motions
Aagaard's figures of different rupture velocities
Spectral representations
Stress drop
Steady-state ruptures do not radiate
Stress in cylindrical hole.
Fracture energy vs yield stress
Residual stress
Energy of a slip pulse.
Frictional energy
Chaotic ruptures
Self-organization
Random walks

Appendix A

A Generalized power-law spectral scaling law

The fact that very near-source accelerations (high-frequencies) appear to be incoherent noise with a peak acceleration that is independent of magnitude motivates the following hypothesis.

Assumption 1: Radiated high-frequency energy $E_R^{f \gg f_c}$ scales with the rupture area, $S = LW$, some unknown power of the stress drop, γ, and is independent of the average slip on that rupture surface, or

$$E_R^{f \gg f_c} \sim S \Delta \sigma^{\gamma} \qquad\qquad 8.210$$

This assumes that high-frequency motions are characterized by random phase, which means that if we sum two motions, then the energies sum linearly. This means that if we double the rupture area, then we double the radiated high-

frequency energy. I have also included a yet unknown static stress drop dependence.

My hypothesis of near-source energies (equation 8.210) and Brune's assumptions are identical if $\gamma = 2$, and if $\sigma_{eff} = \Delta\sigma$. In essence, Brune (1970) also assumed that the near-source high-frequency radiated energy scales linearly with the rupture area (see 8.68).

Assumption 2: Assume that the far-field **seismic wave is approximately a non-negative function of duration** T_c**.** This is equivalent to saying that the motions are approximately the solutions for the far-field S-waves in an elastic whole space. A non-negative function of duration T_c has a flat amplitude spectrum between zero frequency (value is proportional to the integrated area of the function, which is proportional to the Potency) and the corner frequency, $f_c = 1/T_c$. If the S-wave is non negative then its spectrum at periods shorter than T_c is less than the amplitude spectrum for periods longer than T_c . Brune also explicitly made this assumption.

Assumption 3: For simplicity, **assume that the high-frequency spectrum has a power-law high-frequency spectral decay** $f^{-\alpha}$. That is, **assume** that

$$\tilde{U}_R \sim \frac{P}{1 + \left(\frac{f}{f_c}\right)^{\alpha}} \qquad 8.211$$

Where $\tilde{U}_R$ is the Fourier amplitude spectrum of the radiated (far-field terms) S-wave, potency $P = S\bar{D}$, and corner frequency, $f_c = 1/T_c$. Equation 8.211 has asymptotes

$$\tilde{U}_R \sim \begin{cases} P & f \ll f_c \\ P\left(f/f_c\right)^{-\alpha} & f \gg f_c \end{cases} \qquad 8.212$$

Brune also used this assumption, but he assumed that $\alpha = 2$ (based on the analytic solution of an instantaneous shear on a half space).

Assumption 4: Now **assume that the duration of the signal is proportional to the dimension of the fault** (that is, assume constant rupture velocity), or that $f_c \sim \frac{1}{\sqrt{S}}$. Again, Brune used this assumption.

Assuption 5: Assume size similarity of the form $P \sim S\bar{D} \sim S^{3/2}\Delta\sigma$**, or alternatively** $LW \sim P^{2/3}\Delta\sigma^{-2/3}$. Brune also made this assumption. Therefore, $f_c \sim P^{-1/3}\Delta\sigma^{1/3}$, and 8.211 becomes

$$\tilde{U}_R(f) \sim \frac{P}{1 + \dfrac{f^\alpha}{P^{-\alpha/3}\Delta\sigma^{\alpha/3}}} \qquad 8.213$$

Which has asymptotes

$$\tilde{U}_R \sim \begin{cases} P & f \ll f_c \\ P^{1-\alpha/3}\Delta\sigma^{\alpha/3}f^{-\alpha} & f \gg f_c \end{cases} \qquad 8.214$$

Now the radiated energy spectrum scales as the square of the Fourier amplitude spectrum, or at high frequencies

$$E_R^{f \gg f_c} \sim \left(\tilde{U}_R^{f \gg f_c}\right)^2 \sim P^{2(1-\alpha/3)}\Delta\sigma^{2\alpha/3} \qquad 8.215$$

Now my original hypothesis 8.210 was that

$$E_R^{f \gg f_c} \sim S\Delta\sigma^\gamma \sim P^{2/3}\Delta\sigma^{\gamma-2/3} \qquad 8.216$$

The $\Delta\sigma^{-2/3}$ in 8.216 comes from the hypothesis that the radiated high frequencies scale with the rupture area. This hypothesis implies that for the same P , a higher static stress drop has a smaller rupture area and therefore a smaller radiated high-frequency energy.

The scaling of high-frequency radiated energies with P and $\Delta\sigma$ given in the two independent relations 8.215 and 8.216 can be simultaneously true only if $\alpha = 2$ and also if $\gamma = 2$. Assuming $\gamma = 2$ is the same as saying that radiation of high frequency energy per unit of rupture area of scales with the square of the static stress drop. Since energy scales as the square of the wave amplitude, $\beta = 2$ is the same as saying near-source peak acceleration scales linearly with stress drop (this is Brune's assumption). This assumption is essentially the Brune spectral model (with $\Delta\sigma = \sigma_{eff}$). Notice that in this case

$$\tilde{U}_R \sim \frac{P}{1 + \dfrac{f^2}{P^{-2/3}\Delta\sigma^{2/3}}} \sim \begin{cases} P & f \ll f_c \\ P^{1/3}\Delta\sigma^{2/3}f^{-2} & f \gg f_c \end{cases} \qquad 8.217$$

This relation could have been derived from either one of two sets of assumptions. Assumption set 1: the spectrum is an f^{-2} power law with a maximum

amplitude proportional to P and a spectral corner that is inversely proportional to the rupture dimension. Alternatively, assumption set 2: the spectrum is an unknown power law, the high-frequency near-source motions scale linearly with stress drop, and for equal stress drop, the high-frequency radiated energy scales with rupture area (random phase). These two sets of assumptions are equivalent and either set independently leads to the 2nd form of Brune's spectral model.

$$\tilde{U}_R \sim r^3 \Delta\sigma \frac{1}{1 + \dfrac{f^2}{P^{-\frac{2}{3}}\Delta\sigma^{\frac{2}{3}}}}$$

$$\sim \frac{P}{1 + \dfrac{f^2}{P^{-\frac{2}{3}}\Delta\sigma^{\frac{2}{3}}}} \qquad 8.218$$

Which is identical to my earlier statement of Brune's spectral law.

Inconveniently, near-source high-frequencies (pga) seem to be independent of both slip amplitude (see Figure 8.14) and stress drop $\Delta\sigma$ (see Figure 8.11). That is, Figure 8.11 shows that near source pga's are approximately independent of P (for M>6) and $\Delta\sigma$. That is, $\gamma \approx 0$, or rewriting 8.216,

$$E_R^{f \gg f_c} \sim S \sim P^{\frac{2}{3}}\Delta\sigma^{-\frac{2}{3}} \qquad 8.219$$

Which means that the high-frequency spectral amplitude scales as

$$\tilde{U}^{f \gg f_c} \sim \sqrt{S} \sim P^{\frac{1}{3}}\Delta\sigma^{-\frac{1}{3}} \qquad 8.220$$

The lack of correlation between near-source pga and stress drop seen if Fig. 8.11 seems to suggest that the spectral asymptotes are given by

$$\tilde{U}_R \sim \begin{cases} P & f \ll f_c \\ P^{\frac{1}{3}}f^{-2}\Delta\sigma^{-\frac{1}{3}} & f \gg f_c \end{cases} \qquad 8.221$$

Unlike the Brune relationship, we cannot assume a single power law with a single corner to fit the spectrum.

Scale Dependence of the Strength of the Earth's Crust

Thomas H Heaton

Introduction

What is it that earth scientists mean when they speak of the strength of the Earth's crust? Although the concept of strength is well defined for materials that yield uniformly with spatially homogeneous stress, there is little doubt that deformations and stresses are highly heterogeneous in the Earth's crust. I discuss more general definitions of strength that are applicable for materials that have spatially heterogeneous stress when they yield. These new definitions of strength are based on spatial averages of stress, and while they are compatible with a simple intuitive understanding of strength, they lead to the conclusion that materials that are deforming at multiple length scales have strengths that depend on the length scale of the observation. In particular, the strength decreases with increasing length scale.

Strength of a Material

It is natural to think of the strength of a material as the amplitude of stress at which a material begins to yield[2]. In practice, it is common to measure the strength of a material using some test apparatus that applies an increasing external load to a sample until it begins to yield, F_Y^{EXT}. Figure 1 shows a schematic of a hypothetical test apparatus to measure the strength of a cubic sample of dimension L_0. The strength is determined by calculating a yield stress from the yield force and the appropriate cross sectional area L_0^2. For reasons that will become apparent shortly, I will consider the strength to depend on the scale of the sample and I will define it as

$$\Sigma(L_0) \equiv \left| F_Y^{EXT}(L_0) \Big/ L_0^2 \right|. \qquad (8.222)$$

While this is the traditional way that strength is measured for materials, it is not really possible to make this measurement in the Earth's crust; there is no test apparatus large enough to cause the yielding of a sample that is ten's of kilometers in dimension. Of course, one can measure the yield stress of small samples of the crust, but we should be cautious since there may be mechanisms whose deformation physics depend on the scale of the material. Later I will discuss two other definitions of strength, one which is based on root-mean-square statistics, and the other is based on inelastic work.
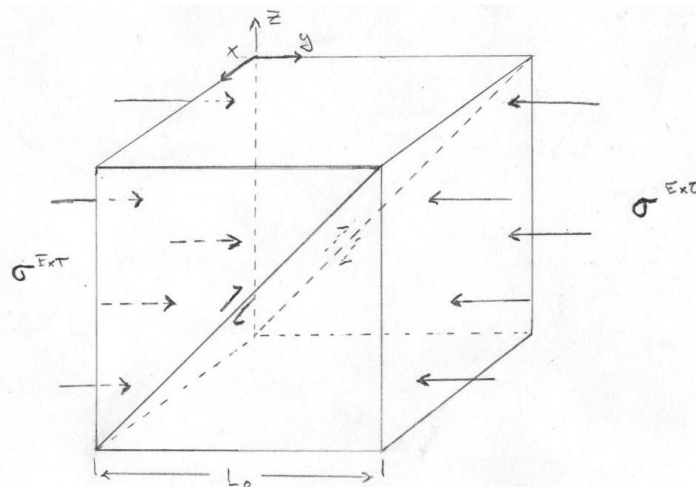

Figure 1. Hypothetical cubic test sample of dimension L_0 that is subjected to a uniform stress. The "strength" of this sample can be viewed as the size of the externally applied stress σ_Y^{Ext} that causes the sample to yield.

Since it is not possible to test the Earth's crust in an apparatus, it is common for earth scientists to estimate the strength of the crust by obtaining estimates of stress amplitudes for a section of the crust that is experiencing inelastic yielding

---

[2] Since stress is a tensor quantity, it is usually necessary to define strength as an imaginary surface in 6-dimensional stress space that separates elastic stress states from yielding stress states. If the material is isotropic, then this yield surface can be defined in the 3-dimensional space of the amplitudes of the principal stresses.

(e.g., earthquakes; see Kanamori and Heaton, 2000). However it is important to recognize that this is a different measurement from the one that is made using a test apparatus in the laboratory. That is, if we return to our conceptual test apparatus of Figure 1, instead of measuring the external yield force F_Y^{EXT} , we would directly measure stress σ_Y inside of our yielding sample. To keep the discussion relatively simple, let's choose to look at only one component of the shear stress in our medium. That is, I could alternatively define the strength of the sample to be the amplitude of internally measured stress within the yielding sample. Although the discussion up to this point may appear to be trivial, the fundamental issue is how to measure this internal yield stress. That is, stress is always spatially variable inside of any solid material. For example, any poly-crystalline material has enormous stresses at grain boundaries; these stresses are on the order of GPa's. Thermal stresses and dislocations are other sources of spatial variations in stress. Therefore, a more complete description of the stress in our medium would be

$$\sigma(x,y,z) = \sigma^{Ext}(L_0) + \sigma^{Int}(x,y,z), \qquad x,y,z \in [0,L_0)$$
$$(8.223)$$

where $\sigma^{Ext}(L_0)$ is the spatially uniform stress from the externally applied force and $\sigma^{Int}(x,y,z)$ is the spatial distribution of internal shear stress.[3] By definition, the spatial average of $\sigma^{Int}(\mathbf{x})$ over the entire sample is zero. However, the condition is stronger than that, the average internal stress over any cross section intersecting the x-axis must also be zero, or

$$\overline{\overline{\sigma^{Int}(\mathbf{x})}}^{L_0} \equiv \frac{1}{L_0^2} \int_0^{L_0} \int_0^{L_0} \sigma^{Int}(x,y,z)\, dydz = 0, \qquad (8.224)$$

where I use the double bar notation to signify the average over a 2-dimensional surface that is perpendicular to the x-axis. Therefore, by using this decomposition, I can alternatively define the strength as

$$\Sigma(L_0) \equiv \left| \overline{\overline{\sigma(\mathbf{x})}}^{L_0} \right| \equiv \left| \frac{1}{L_0^2} \int_0^{L_0} \int_0^{L_0} \sigma(x,y,z)\, dydz \right| = \left| \sigma_Y^{Ext}(L_0) \right|. \quad (8.225)$$

Now suppose that at a given external load, then the sample could yield at either the scale length of the sample, or it could also yield only in some smaller section of our sample with the dimension, L_1. Therefore, we would conclude that the strength of the smaller cube is

$$\Sigma(L_1) \equiv \left| \overline{\overline{\sigma(\mathbf{x})}}^{L_1} \right| \equiv \left| \frac{1}{L_1^2} \int_0^{L_1} \int_0^{L_1} \sigma(x,y,z)\, dydz \right| = \left| \Sigma(L_0) + \frac{1}{L_1^2} \int_0^{L_1} \int_0^{L_1} \sigma^{Int}(x,y,z)\, dydz \right|$$
$$(8.226)$$

---

[3] Internal stresses are sometimes called "residual stress", although we will use a more general description of these stresses than is usually considered.

Therefore, $\Sigma(L_1) \neq \Sigma(L_0)$ unless $\overline{\overline{\sigma^{Int}(x,y,z)}}^{L_1} = 0$. Importantly, earthquakes seem to happen at multiple length scales $L_i, i=1...n$, which implies that the stress in the crust is such that the crust is at its yield strength at multiple length scales. That is, the strength of the crust is independent of the length scale if, and only if, $\overline{\overline{\sigma^{Int}(x,y,z)}}^{L_i} = 0$ for all length scales L_i. If $i \to \infty$, then for the strength to be independent of scale length, then $\sigma^{Int}(\mathbf{x}) = 0$, which would imply homogeneous stress in the crust, which is inconsistent with the occurrence of earthquakes at multiple length scales. Therefore, if the crust is failing at all length scales, then the strength of the crust, as defined by (8.225), must depend on the scale length.

More generally, since the definitions of strength given by (8.225) and (8.222) are equivalent, this implies that if our laboratory apparatus recreates the same physics as earthquake failures in the crust, then we would also measure changes in the strength of samples as a function of the size of the sample. As we will see, the key to a deeper understanding of this problem is to characterize the internal stress distribution in the crust.

A Statistical Description of Stress in the Crust

In the previous section I defined the strength of a yielding volume as the amplitude of the average stress on a plane that cuts through the volume, that is, equation(8.225). I am particularly interested in how strength Σ varies as a function of the length scale L of the yielding volume. In the case of the Earth's crust, I **make the following key assumption. The crust has evolved into such a state of stress that it can fail at any length scale.** That is, I assume that earthquakes of any size are possible within our hypothetical crust (Bak and others, 1987). In this case the distribution of stress $\sigma(x,y,z)$ tells us the length scale dependence of the strength in the crust. All we need to do is to estimate the amplitude of the spatial average of $\sigma(x,y,z)$ as a function of the length scale.

At this point, it is convenient to assume that $\sigma(x,y,z)$ is a random stationary function of y and z. The assumption of stationarity means the statistical properties of the stress are invariant with respect to position within the medium. That is, the joint statistical distribution of $\sigma^{Int}(x_1,y_1,z_1),...,\sigma^{Int}(x_n,y_n,z_n)$ is the same as that of $\sigma^{Int}(x_1+x,y_1+y,z_1+z),...,\sigma^{Int}(x_n+x,y_n+y,z_n+z)$. Assuming stationarity means that the process σ^{Int} is entirely determined by its covariance function

$$\Gamma(x,y,z,x',y',z') \equiv \text{cov}\left[\sigma^{Int}(x,y,z), \sigma^{Int}(x',y',z')\right]$$
$$\equiv E\left\{\left[\sigma^{Int}(x,y,z) - E\left(\sigma^{Int}(x,y,z)\right)\right]\left[\sigma^{Int}(x',y',z') - E\left(\sigma^{Int}(x',y',z')\right)\right]\right\}$$
$$(8.227)$$

which is also known under the names of the autocorrelation function or the power spectrum, and where E is taken to mean *expected value*. Note that the stationarity assumption implies that the covariance between two points depends

only on the distance between those points. Therefore, we can define

$$\gamma\left(\xi - \xi'\right) \equiv \Gamma\left(\xi, \xi'\right)$$

I recognize that stress in the Earth's crust is clearly not stationary with respect to position; for example, there are systematic variations in crustal thickness, material properties, total deformation, and deformation rate. Nevertheless, this assumption is necessary for later analysis in this paper and it may be approximately valid within regions of the crust. Furthermore, if I choose an x-axis that is perpendicular to a fault plane, then I am essentially assuming that the statistical properties of stress on the fault plane are independent of the position on the fault plane.

I can now restate my definition of strength in (8.225) as

$$\Sigma(L) \equiv E\left(\left\|\overline{\overline{\sigma\left(x,y,z\right)}}^{L}\right\|\right) \tag{8.228}$$

Where E is the expected value of the amplitude of the stress averaged over the volume of length scale L. As an alternative, it can make the mathematics simpler if I introduce an alternative definition of strength that I will call the root-mean-square strength $\Sigma_{rms}(L)$ and which I define as

$$\Sigma_{rms}^2(L) \equiv E\left[\left(\overline{\overline{\sigma\left(x,y,z\right)}}^{L}\right)^2\right] \tag{8.229}$$

Using rms in the definition ensures that strength is always a positive amplitude, but it has the advantage that it allows me to relate strength to the statistical variance of stress. In particular, I show in Appendix I (see page 316 of Dwass) that

$$\Sigma_{rms}^2(L) \equiv \left[E\left(\overline{\overline{\sigma\left(x,y,z\right)}}^{L}\right)\right]^2 + Var\left(\overline{\overline{\sigma\left(x,y,z\right)}}^{L}\right)$$
$$= \left(\sigma_Y^{Ext}(L_0)\right)^2 + Var\left(\overline{\overline{\sigma^{Int}\left(x,y,z\right)}}^{L}\right) \tag{8.230}$$

This means that I can find the length scale dependence of *rms* strength by finding the dependence of the variance of the spatially averaged stress on the length scale of the spatial averaging.

I can use Parseval's theorem to calculate the variance of a stationary function that is described by its covariance. That is, the integrated energy in the space domain, which is the variance times the length of the function, is identically equal to the integral of the power spectrum at all frequencies. That is, if $\overline{\overline{\tilde{\sigma}^{Int}}}\left(k_x, k_y, k_z\right)$ is the 3-dimensional Fourier transform of the spatially averaged internal stress, then it has a variance given by

$$Var\left(\overline{\overline{\sigma^{Int}(x,y,z)}}^{L}\right) = \frac{1}{L_0^2} \int_0^{L_0} \int_0^{L_0} \left|\overline{\overline{\sigma(x,y,z)^{Int}}}^{L}\right|^2 dydz$$

$$= \frac{1}{L_0^2} \int_{L_0^{-1}}^{\infty} \int_{L_0^{-1}}^{\infty} \left|\overline{\overline{\tilde{\sigma}(k_x,k_y,k_z)^{Int}}}^{L}\right|^2 dk_y dk_z$$

(8.231)

If I can characterize how the integral in (8.231) depends on scale length L, then I will have solved our problem. At this point, I assume that $\sigma^{Int}(x,y,z)$ is a mean-zero (i.e., $E\left[\sigma^{Int}(x,y,z)=0\right]$) Gaussian stationary process. That is, I assume that the internal stress can be approximated by a spatially filtered version of Gaussian white noise. I am unaware of any physical basis for this assumption, and I use it because it makes the mathematics simpler.

I am particularly interested in the obtaining the statistical properties of $\overline{\overline{\sigma^{Int}(x,y,z)}}^{L}$, which is the spatial average of the stress over a 2-dimensional surfaces of dimension L . Obtaining the spatial average of the stress is equivalent to 2-dimensional convolution of the stress with some 2-dimensional function $g\left(\frac{y}{L},\frac{z}{L}\right)L^{-2}$, where g has circular symmetry with respect to y and z, and g has unit characteristic length and integrated area (e.g., a Gaussian, a box function, etc.). That is, $g\left(\frac{y}{L},\frac{z}{L}\right)L^{-2}$ has a characteristic length of L and an integrated area of unity. Therefore

$$\overline{\overline{\sigma^{Int}(x,y,z)}}^{L} = L^{-2}\sigma^{Int}(x,y,z)*g\left(\frac{y}{L},\frac{z}{L}\right)$$

$$= L^{-2}\sigma^{Int}(x,y,z)*\left[g'\left(\frac{y}{L}\right)\delta(z)\right]*\left[g'\left(\frac{z}{L}\right)\delta(y)\right]$$

(8.232)

Where g' is the 1-dimensional version of g. Therefore I can conclude that

$$\overline{\overline{\tilde{\sigma}^{Int}(k_x,k_y,k_z)}}^{L} = L^{-2}\tilde{\sigma}^{Int}(k_x,k_y,k_z)\tilde{g}'(Lk_y)\tilde{g}'(Lk_z) \quad (8.233)$$

I can substitute (8.233) into (8.231) to conclude that

$$Var\left(\overline{\overline{\sigma^{Int}(x,y,z)}}^{L}\right) = \frac{1}{L_0^2} \int_{L_0^{-1}}^{\infty} \int_{L_0^{-1}}^{\infty} \left|\tilde{\sigma}^{Int}(k_x,k_y,k_z)\tilde{g}'(Lk_y)\tilde{g}'(Lk_z)\right|^2 dk_y dk_z$$

(8.234)

At this point, it is helpful to be more specific about g'. If I were to assume g' to be a box rectangle function in the space domain (corresponding to an un-weighted running mean), then it's transform would be a sinc function in the wavenumber domain. If, on the other hand, I assume that g' is a sinc function in the space domain, then its transform is a rectangle function in the wavenumber domain and the solution becomes particularly simple. That is if

$$g'(y) \equiv \frac{\sin(y)}{y} \qquad (8.235)$$

Then

$$\tilde{g}'(k_y) = \Pi(k_y) \equiv \begin{cases} 1, & -\frac{1}{2} < k_y < \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \qquad (8.236)$$

Substituting (8.236) into (8.234), I obtain

$$Var\left(\overline{\overline{\sigma^{Int}(x,y,z)}}^L\right) = \frac{1}{L_0^2} \int_{L_0^{-1}}^{\infty} \int_{L_0^{-1}}^{\infty} \left|\tilde{\sigma}^{Int}\left(k_x,k_y,k_z\right)\Pi\left(Lk_y\right)\Pi\left(Lk_z\right)\right|^2 dk_y dk_z$$

$$= \frac{1}{L_0^2} \int_{L_0^{-1}}^{L^{-1}} \int_{L_0^{-1}}^{L^{-1}} \left|\tilde{\sigma}^{Int}\left(k_x,k_y,k_z\right)\right|^2 dk_y dk_z$$

$$(8.237)$$

Substituting (8.237) into (8.230) gives the functional dependence of *rms* strength on length scale.

$$\Sigma_{rms}^2(L) = \left(\sigma_Y^{Ext}(L_0)\right)^2 + \frac{1}{L_0^2} \int_{L_0^{-1}}^{L^{-1}} \int_{L_0^{-1}}^{L^{-1}} \left|\tilde{\sigma}^{Int}\left(k_x,k_y,k_z\right)\right|^2 dk_y dk_z \quad (8.238)$$

This relationship tells us how to determine the strength of the sample at smaller scales L than the scale L_0 of the entire sample. Conversely, if I already knew the rms strength of the material at a smaller scale L, then I could determine its value at the larger scale L_0 simply by rearranging (8.238) as

$$\left(\sigma_Y^{Ext}(L_0)\right)^2 = \Sigma_{rms}^2(L) - \frac{1}{L_0^2} \int_{L_0^{-1}}^{L^{-1}} \int_{L_0^{-1}}^{L^{-1}} \left|\tilde{\sigma}^{Int}\left(k_x,k_y,k_z\right)\right|^2 dk_y dk_z \quad (8.239)$$

That is, if I could make in situ measurements of stress in the crust at the scale length of meters, then I could infer the *rms* strength at the length scale of kilometers by using (8.239), provided of course that I can characterize $\left|\tilde{\sigma}^{Int}\left(k_x,k_y,k_z\right)\right|$.

The nature of stress heterogeneity in the crust

.As mentioned in the introduction, we can view the question of strength of the crust as being equivalent to characterizing $\left|\tilde{\sigma}\left(k_x,k_y,k_z\right)\right|$. That is, through millennia of deformation, the crust has evolved into a state of stress whose statistical characteristics are determined by its spatial power spectrum. In a very fundamental way, knowing this power spectrum tells us the strength of the crust. However, a power spectral definition of strength is definitely not as simple as the more traditional notion of measuring a force necessary to yield a sample (i.e.,

equation (8.222)). We saw that these two notions of strength can be connected through equation (8.238).

What is the nature of $\left|\tilde{\sigma}\left(k_x, k_y, k_z\right)\right|$? To keep things as simple as possible, we will assume that the power spectrum is isotropic in space. Although we have explicitly acknowledged the 3-dimensionality of the crust, we have adopted a definition of strength that is based on the failure of 2-dimensional surfaces. In some ways it is attractive to drop one of the dimensions from this discussion. However, we will later discuss seismicity data that is inherently 3-dimensional. If the stress power spectrum is isotropic in three dimensions, then we can consider that we can write it as $\left|\tilde{\sigma}\left(k_r\right)\right|$, where $k_r = \sqrt{k_x^2 + k_y^2 + k_z^2}$.

If the stress in the crust was a simply a constant (which it's clearly not), then the strength is also a constant, independent of the scale. In this case, $\left|\tilde{\sigma}^{Int}\left(k_r\right)\right| = 0$ and the power spectrum of stress is simply a 3-dimensional impulse function at the origin, or $\left|\tilde{\sigma}\left(k_r\right)\right| = \Sigma\left(L_0\right)\delta\left(k_r\right)$.

A more interesting case is to investigate the possibility that the power spectrum can be approximated as a power law, or assume that $\left|\tilde{\sigma}\left(k_r\right)\right| = Ck_r^{-\beta}$, where β is positive and where the outer scale of the sample is infinite. In this case, the strength approaches zero as the scale becomes infinite and (8.238) can be written as

$$\Sigma_{rms}^2\left(L\right) = \frac{C}{L_0^2} \int_0^{L^{-1}} \int_0^{L^{-1}} \left(k_y^2 + k_z^2\right)^{-\beta} dk_y dk_z \qquad (8.240)$$

This double integral over a rectangular box in Cartesian coordinates can be transformed to a double integral over a circular area without any loss of generality. That is, we can rewrite (8.240) as

$$\Sigma_{rms}^2\left(L\right) = C' \int_0^{L^{-1}} \int_0^{2\pi} k_r^{-2\beta} \left(k_r d\theta\right) dk_r$$

$$= 2\pi C' \int_0^{L^{-1}} k_r^{1-2\beta} dk_r \qquad (8.241)$$

$$= \frac{2\pi C'}{\left(1-2\beta\right)} \frac{1}{L^{2(1-\beta)}}$$

Or

$$\Sigma_{rms}\left(L\right) = \frac{C''}{\sqrt{1-2\beta}} \frac{1}{L^{1-\beta}}, \quad 0 \le \beta < 1 \qquad (8.242)$$

The remarkably simple result of (8.242) only applies if the material is in a critical state at all length scales. Or, in other words, the material could experience an event of any size. The spectral decay parameter β must be smaller than 1, since if it is larger than 1, it would lead to the conclusion that long wavelength variations in stress are large compared to small ones. Such a stress state is incompatible with the concept that failure can occur at any length scale.

In many ways, either uniform stress, or random stress with a power law spectral content, are the stress states requiring the fewest parameters to describe them. Both of these idealized stress states have no inherent length scale. Of course, the stress in the Earth is determined by failure processes that introduce scale lengths that result in power spectra that are not easily parameterized. However, we will discuss evidence that suggests that earthquake failure processes may be consistent with power-law spectral behavior over a relatively broad range of scale lengths.

Numerical simulations using fractal tensors

Up to this point we have discussed stress as if it were a scalar quantity, which is clearly not the case. Smith and Heaton (2006) demonstrate a procedure to generate spatially varying stochastic stress tensors that 1) are isotropic (no preferred orientations), 2) has scalar invariants that are power law for any section through the material, and 3) has orientations of principal coordinate frames that vary stochastically in space with correlation at all length scales (i.e., the orientations are fractal). The procedure consists of parameterizing the stress at each point with three scalar invariants and with the orientation of the principal coordinate frame (another 3 degrees of freedom) for each point. For the purpose of this calculation, we are interested in spatial averages of shear stress, so we can assume that the scalar invariant corresponding to pressure is spatially constant (that is, it drops out of the problem).

We start with a cubic grid of discrete points. We then assign zero-mean Gaussian random numbers corresponding to two principal stresses. The third principal stress is then chosen so that the sum of the principal stresses is zero. These stress invariants are then spatially filtered in three dimensions using a wavenumber filter given by $\tilde{f}_{ijk} = n^{-3} \left(i^2 + j^2 + k^2 \right)^{-\frac{1}{2}\beta}$, $i, j, k = 1...n$, where n is the number of points along a linear axis.

Thin faults lead to fractal stress because there is no scale length

Strong velocity weakening leads to very unstable slip

Slip length scaling can be approximated with a self-similar statistical models

Focal mechanism statistics

Radiated energy may be controlled by prestress

Kanamori, H., and T. Heaton, 2000, Microscopic and macroscopic physics of earthquakes, contained in Geocomplexity and the Physics of Earthquakes, Editors J. Rundle, D. Turcotte, and W. Klein, Geophysical Monograph 20, Published by the American Geophysical Union, D.C., 127-141.