Earth Planets Space, **56**, 773–793, 2004

Earthquake cycles and physical modeling of the process leading up to a large earthquake

Mitiyasu Ohnaka*

The University of Tokyo, and University College London

(Received November 30, 2003; Revised June 15, 2004; Accepted July 1, 2004)

A thorough discussion is made on what the rational constitutive law for earthquake ruptures ought to be from the standpoint of the physics of rock friction and fracture on the basis of solid facts observed in the laboratory. From this standpoint, it is concluded that the constitutive law should be a slip-dependent law with parameters that may depend on slip rate or time. With the long-term goal of establishing a rational methodology of forecasting large earthquakes, the entire process of one cycle for a typical, large earthquake is modeled, and a comprehensive scenario that unifies individual models for intermediate- and short-term (immediate) forecasts is presented within the framework based on the slip-dependent constitutive law and the earthquake cycle model. The earthquake cycle includes the phase of accumulation of elastic strain energy with tectonic loading (phase II), and the phase of rupture nucleation at the critical stage where an adequate amount of the elastic strain energy has been stored (phase III). Phase II plays a critical role in physical modeling of intermediate-term forecasting, and phase III in physical modeling of short-term (immediate) forecasting. The seismogenic layer and individual faults therein are inhomogeneous, and some of the physical quantities inherent in earthquake ruptures exhibit scale-dependence. It is therefore critically important to incorporate the properties of inhomogeneity and physical scaling, in order to construct realistic, unified scenarios with predictive capability. The scenario presented may be significant and useful as a necessary first step for establishing the methodology for forecasting large earthquakes.

Key words: Rational constitutive law, earthquake rupture, physics of rock friction and fracture, inhomogeniety, physical scaling, physical modeling, a unified scenario for earthquake forecasting.

1. Introduction

Two important properties have to be considered for realistic modeling of the earthquake generation process: that is, inhomogeneity, and physical scaling. The seismogenic layer and individual faults therein are inherently inhomogeneous. It is widely recognized that a large earthquake at shallow crustal depths never occurs alone, but is necessarily accompanied by aftershocks, and often preceded by seismic activity (small to moderate earthquakes) enhanced in a relatively wide region surrounding the fault during the process leading up to the event. This is a reflection of the above fact that the seismogenic layer is inhomogeneous.

If, for instance, an entire earthquake fault were homogeneous and very weak, strong motion seismic waves would not be generated. For strong motion waves to be generated, the fault itself must be inhomogeneous, and include local areas (which may be called patches) of high resistance to rupture growth in the fault zone. Indeed, seismological observations and their analyses (e.g., Kanamori and Stewart, 1978; Aki, 1979, 1984; Kanamori, 1981; Bouchon, 1997) commonly show that individual faults are heterogeneous, and include what is called "asperities" or "barriers". The presence of "asperities" or "barriers" on a fault is a clear manifestation that real faults comprise strong portions of high resistance to rupture growth with the rest of the fault having low (or little) resistance to rupture growth.

The resistance to rupture growth has a specific physical meaning in the framework of fracture mechanics, and it is defined as the shear rupture energy required for the rupture front to further grow (see Ohnaka, 2000). It has been suggested from laboratory experiments that some of the "asperities" or "barriers" on an earthquake fault are strong enough to equal the strength of intact rock (Ohnaka, 2003). Such strong portions of high resistance to rupture growth on a fault are required for an adequate amount of the elastic strain energy to accumulate in the elastic medium surrounding the fault with tectonic loading, as a driving force to bring about a large earthquake or to radiate strong motion seismic waves.

We can thus conclude that an earthquake rupture at shallow crustal depths is shear rupture instability that takes place on an inhomogeneous fault embedded in the seismogenic layer, which is also inhomogeneous. Fault inhomogeneity includes geometric irregularity of the rupture surfaces, which in turn not only causes stress inhomogeneity but also plays an important role in scaling scale-dependent physical quantities inherent in the rupture (e.g., Ohnaka and Shen, 1999; Ohnaka, 2003). More specifically, it has been found with recent laboratory experiments (Ohnaka, 2003) that the fundamental cause of the scaling property lies at the characteristic length scale defined as the predominant wavelength repre-

*Present address: Utsukushigaoka-nishi 3-40-19, Aoba-ku, Yokohama 225-0001, Japan.

senting geometric irregularity of the rupture surfaces.

Thus, the properties of inhomogeneity and physical scaling are the key to physical modeling of the process leading up to a large earthquake, and hence they must be incorporated into its physical model. The purpose of this paper is first to discuss thoroughly what the rational constitutive law for earthquake ruptures ought to be from the standpoint of rock physics on the basis of solid facts observed in the laboratory. We will then present a model of the cyclic process for a typical, large earthquake. Finally, we will show how consistently the process leading up to a typical, large earthquake can be modeled in terms of the underlying physics within the context of the earthquake cycle model, by incorporating the properties of inhomogeneity and physical scaling. The models presented here will be significant and useful as a necessary step for establishing the methodology for forecasting large earthquakes.

2. Constitutive Formulation for Earthquake Ruptures

It has been established to date that the shear rupture process is governed by the constitutive law. However, it is still controversial how the constitutive law for earthquake ruptures should be formulated. It is therefore critically important to discuss closely what it ought to be. I wish to discuss in this section how the constitutive law for earthquake ruptures should be formulated from the standpoint of the physics of rock friction and fracture on the basis of solid evidence observed in the laboratory. The constitutive formulations so far attempted can be categorized into two different approaches: the rate-dependent formulation, and the slip-dependent formulation.

One of the simplest attempts to formulate the constitutive law for earthquake ruptures may be to assume that the shear traction τ is a function of slip rate (or velocity) $\dot{D}$ alone, which is characterized by a velocity-weakening property (e.g., Carlson and Langer, 1989; Carlson *et al.*, 1991; Nakanishi, 1992). This formulation, however, does not lead to a self-consistent constitutive law (Rice and Ruina, 1983; Ruina, 1985), because it predicts that the rupture is necessarily unstable since $d\tau/d\dot{D} < 0$. This contradicts the common observation that the shear rupture can proceed stably even in the purely brittle regime. The formulation also contradicts the observational fact that the shear traction during the dynamic process is not a single-valued function of the slip rate (Ohnaka *et al.*, 1987, 1997; Ohnaka and Yamashita, 1989). To resolve these contradictions, Dieterich (1978, 1979, 1981, 1986) and Ruina (1983, 1985) introduced an evolving state variable which is a measure of the quality of surface contact, and they proposed a rate- and state-dependent constitutive law for frictional slip failure.

The rate- and state-dependent constitutive formulation assumes that the slip rate $\dot{D}$ and, at least, one evolving state variable Θ are independent and fundamental variables, and that the transient response of the shear traction τ to $\dot{D}$ is essentially important. In this formulation, therefore, the role of $\dot{D}$ is emphasized, and τ is expressed as an explicit function of $\dot{D}$ and Θ. This formulation is based on experimental data observed at very slow slip speeds less than 1 mm/s (Dieterich, 1978, 1981), and on their interpretation on those data. The law can specifically be expressed as (Dieterich, 1986; Okubo, 1989; Linker and Dieterich, 1992):

$$\tau = (\mu_0 + \mu_1 - \mu_2)\sigma_n^{\text{eff}} \tag{1a}$$

$$\frac{d\Theta}{dt} = 1 - \frac{\Theta \dot{D}}{d_c} \tag{1b}$$

where

$$\mu_1 = b \ln\left(\frac{\Theta \dot{D}_*}{d_c} + 1\right) \tag{1c}$$

and

$$\mu_2 = a \ln\left(\frac{\dot{D}_*}{\dot{D}} + 1\right). \tag{1d}$$

In the above equations, σ_n^{eff} represents the effective normal stress defined as $\sigma_n^{\text{eff}} = \sigma_n - P_p$ (σ_n, normal stress; P_p, pore fluid pressure), μ_0 represents frictional coefficient independent of Θ and $\dot{D}$, μ_1 represents the contribution of state variable Θ to friction, μ_2 represents the contribution of slip rate $\dot{D}$ to friction, $\dot{D}_*$ represents the reference slip rate, and a, b and d_c are the constitutive law parameters. The above equations have been derived by considering the effects of Θ and $\dot{D}$ alone, under the condition that direct effect of slip displacement D on friction is negligible ($\partial\mu/\partial D \cong 0$), which may be attained only after an adequate amount of the slip displacement. The assumption that μ_0 is constant is justifiable only under the condition that $\partial\mu/\partial D = 0$. Note, however, that the direct effect of slip displacement may be more dominant during actual rupture processes than the effects of Θ and $\dot{D}$, which will specifically be discussed later in this section.

Using a set of the above equations, Bizzarri *et al.* (2001), and Cocco and Bizzarri (2002) performed two-dimensional numerical simulations of dynamic rupture regime at high slip speeds. Based on their simulated results, they argued that "there is no need to assume that friction must become independent of slip rate at high speeds to resemble slip weakening" (Cocco and Bizzarri, 2002). However, this argument seems logically inconsistent, because the effect of high velocity cutoff has been incorporated into Eq. (1d) used in their simulation. Note that the cutoffs to the rate- and state-dependencies have been made by including the $+1$ term in the argument of the logarithm in Eqs. (1c) and (1d) (Okubo and Dieterich, 1986; Dieterich, 1986; Okubo, 1989). In particular, Eq. (1d) is formulated so as for the effect of $\dot{D}$ on friction to saturate at high slip rates, and this saturation (or high velocity cutoff) is required for agreeing with the facts observed in the dynamic regime of frictional slip (Okubo and Dieterich, 1986; Okubo, 1989). Hence, the parameter $\dot{D}_*$ cannot have an arbitrary value, but has to have a specific value constrained by the observation. Bizzarri *et al.* (2001) concluded in their paper that the rate- and state-dependent constitutive formulation yields a complete description of the rupture process. However, their computer-simulated results are not compared with any observational evidence, and therefore it is not clear from their paper how and to what extent their simulated results quantitatively account for dynamic regime of rupture processes in the real world. For the physics of earthquakes to aim for an exact science, I believe it is critically important to check in quantitative terms how simulated results agree with the observational evidence.

From the fact that the effect of $\dot{D}$ on frictional sliding saturates at high slip velocities, it follows that the dynamic rupture regime at high slip velocities is independent of $\dot{D}$. This indicates that $\dot{D}$ is not appropriate as an independent variable for the constitutive formulation, at least in the dynamic regime of high slip velocities. In addition, this particular formulation does not lead to a unifying constitutive law that governs both frictional slip failure and the shear fracture of intact rock, which is a prerequisite for the constitutive formulation for earthquake ruptures (Ohnaka, 2003). Hence, the rate-dependent formulation cannot be the governing law for earthquake ruptures that are a mixture of frictional slip failure and the shear fracture of intact rock.

Although the rate effect has been found for frictional sliding of wide materials (Dieterich and Kilgore, 1996), its quantitative effect is very small if compared with the effects of the displacement and the effective normal stress. It should be noted that the rate effect can be measured only after an adequate amount of slip displacement by which the direct effect of slip displacement (or $\partial\mu/\partial D$) has been reduced (Dieterich, 1979, 1981). Indeed, laboratory experiments show that the parameter a has a very small value ranging from 0.003 to 0.015 for granite rock (Dieterich, 1981; Gu *et al.*, 1984; Tullis and Weeks, 1986; Blanpied *et al.*, 1987). Thus, the effect of slip rate during the rupture process may be masked completely by a dominant effect induced by the slip displacement on friction on an inhomogeneous fault whose surfaces are geometrically irregular. The rate effect may also be masked by the effect of perturbation (or fluctuation) of the normal stress and/or pore water pressure, which possibly occurs in the fault zone during the rupture process. Indeed, it has been shown that the rate effect is not necessarily required for explaining dynamic rupture regime of actual earthquakes if their fault inhomogeneities, which are an inherent property of natural faults in the Earth's crust, are taken into account (Beroza and Mikumo, 1996). Hence, I believe that there is no compelling reason to emphasize the effect of the slip rate in the constitutive formulation for earthquake ruptures.

One may argue that the rate effect is more important in a much slower slip rate range. Indeed, laboratory experiments show that frictional resistance increases linearly with a logarithmic decrease in the slip rate in the range of very slow rates (Dieterich, 1978). This effect may certainly play a significant role in fault re-strengthening (or healing) after the arrest of dynamic rupture. One must recognize, however, that the slip rate effect is not the only mechanism for fault re-strengthening. There are other mechanisms for the fault re-strengthening. Aochi and Matsu'ura (2002) showed that the fault re-strengthening can be attained if the time-dependence of adhesion of real contact areas due to such an effect as thermal diffusion is considered on the fault surfaces. Real rupture surfaces of inhomogeneous rock are not flat planes, but exhibit geometric irregularity. The re-strengthening on such irregular fault surfaces can also be attained by a displacement-induced increase in friction, due to an increase in the sum of the real areas of asperity contact, interlocking, and/or ploughing on the fault with proceeding displacement. Although this direct effect of displacement has been overlooked in the constitutive formulation, I emphasize that it can in reality be a dominant mechanism for the time-dependent increase in frictional resistance on the fault under the compressive stress. The re-strengthening may also be reinforced by a gradual increase in the effective normal stress with tectonic loading during the inter-seismic period. Thus, the fault re-strengthening can easily be attained without having to assume the effect of slip rate, so that we again cannot find any compelling reason to emphasize the slip rate effect in the constitutive formulation for earthquake ruptures.

As understood from the basic fact that three fundamental modes (mode I, II, and III) of fracture are defined in terms of the crack-tip displacement in fracture mechanics, the displacement plays a fundamental and primary role in the fracturing process. One has to recognize that the slip-dependency is a more fundamental property of the shear rupture than the rate-dependency, and this basic fact must be taken into account when the constitutive law for earthquake ruptures is formulated. Thus, the governing law for earthquake ruptures should be formulated in such a manner that the shear traction τ is a primary function of the slip displacement D, with its functional form that may be affected by a parameter of slip rate $\dot{D}$ or stationary contact time. This formulation assumes that the slip displacement is an independent and fundamental variable, and that the transient response of the shear traction to the slip displacement is essentially important.

The slip-dependent constitutive relation derived from laboratory experiments can commonly be illustrated as shown in Fig. 1(a). This constitutive relation is self-consistent as the governing law for the shear rupture (Rice, 1980, 1983, 1984; Rudnicki, 1980, 1988). In addition, the slip-dependent formulation not only makes it possible to unify both frictional slip failure on a pre-existing fault and the shear fracture of intact rock consistently (Ohnaka, 2003), but also is justified from the standpoint of microcontact physics (Matsu'ura *et al.*, 1992). Hence, it is quite reasonable from physical viewpoints to assume the slip-dependent constitutive law as the governing law for earthquake ruptures.

The constitutive law may be expressed as (Ohnaka, 1996; Ohnaka *et al.*, 1997):

$$\tau = f(D; \dot{D}, \lambda_c, \sigma_n^{\mathrm{eff}}, T, \mathrm{CE}) \qquad (2)$$

where f represents the constitutive relation between τ and D, which is in general affected by not only $\dot{D}$ but also such parameters as scaling parameter λ_c (which will be defined later), σ_n^{eff}, temperature T, and chemical effect of interstitial pore water CE. We have to keep in mind that the effect of $\dot{D}$ is secondary compared with the primary effect of D (this is the reason why the law is referred to as *slip-dependent* law).

A simplified model (see Fig. 1(b)) of the functional form (2) has often been employed for theoretical modeling of earthquake ruptures and for their numerical simulations (Ida, 1972; Andrews, 1976a, b; Day, 1982; Campillo and Ionescu, 1997; Campillo *et al.*, 2001; Madariaga *et al.*, 1998; Madariaga and Olsen, 2000, 2002; Bizzarri *et al.*, 2001; Fukuyama and Madariaga, 2000; Fukuyama and Olsen, 2002; Uenishi and Rice, 2003). Although such a simplified slip-weakening model is certainly useful, note that the model necessarily leads to a singularity of slip acceleration near the front of a dynamically propagating rupture (Ida, 1973), which is physically unrealistic. To avoid such an un-

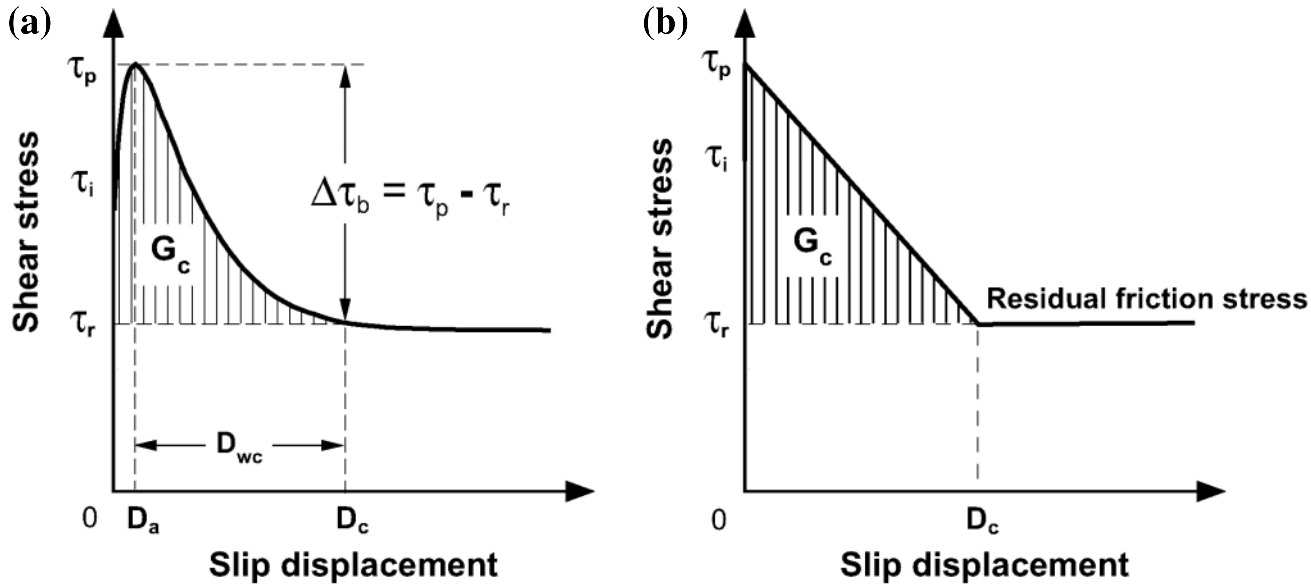

Fig. 1. (a) A slip-dependent constitutive relation for the shear rupture. In the figure, τ_i is the initial strength on the verge of slip, τ_p is the peak shear strength, $\Delta\tau_b$ is the breakdown stress drop defined by $\Delta\tau_b = \tau_p - \tau_r$ (τ_r, residual frictional stress), D_a is the critical displacement at which the peak shear strength is attained, and D_c is the breakdown displacement defined as the critical slip required for the shear traction to degrade to the residual frictional stress. The shear rupture energy G_c is equal to the area of the hatched portion. (b) A simplified slip-weakening constitutive relation. In this model, no slip displacement is necessary for the shear stress to increase from the initial strength τ_i to the peak shear strength τ_p. D_c is the breakdown displacement required for the shear traction to degrade to the residual frictional stress τ_r, and G_c is the shear rupture energy.

realistic singularity of slip acceleration, specific functional form (2) has to be determined so as to incorporate the slip-strengthening property (Ida, 1973; Ohnaka and Yamashita, 1989). This is particularly important when strong motion source parameters such as the peak slip velocity and acceleration in dynamic rupture regime of high slip velocities are discussed in quantitative terms (Ohnaka and Yamashita, 1989). Specific expressions of the form (2) have been proposed by earlier authors (Ohnaka and Yamashita, 1989; Matsu'ura *et al.*, 1992; Ohnaka, 1996; Aochi and Matsu'ura, 2002); however, its mathematical expression is not of primary concern, so that it will not be presented here.

The effects of $\dot{D}$, λ_c, σ_n^{eff}, T, and CE are implicitly exerted on the law through the constitutive law parameters τ_i, τ_p, $\Delta\tau_b$, D_a, and D_c. Here, τ_i is the initial shear stress on the verge of slip, τ_p is the peak shear strength, $\Delta\tau_b$ is the breakdown stress drop defined as $\Delta\tau_b = \tau_p - \tau_r$ (τ_r, residual friction stress), D_a is the critical slip at which the peak strength is attained, and D_c is the breakdown slip displacement defined as the critical slip required for the shear traction to degrade to τ_r. For instance, laboratory experiments show that τ_p depends on $\dot{D}$, σ_n^{eff}, T, and CE, and hence τ_p may in general be written as (Ohnaka *et al.*, 1997):

$$\tau_p = \tau_p(\dot{D}, \sigma_n^{\mathrm{eff}}, T, \mathrm{CE}) \quad (3)$$

It has been documented that there are two competing rate effects on τ_p, one of which has already been discussed (see Eq. (1d)). The rate effect formulated by Eq. (1d) is operative during frictional sliding on a pre-existing fault. The other rate effect has been found for the shear fracture of intact rock (Masuda *et al.*, 1987; Kato *et al.*, 2003b). This rate effect is more enhanced in wet environments than in dry environments (Kato *et al.*, 2003b), so that the mechanism may be ascribed to stress-aided corrosion. Its quantitative effect on τ_p also obeys a logarithmic law (Masuda *et al.*, 1987; Kato *et al.*, 2003b); that is, τ_p decreases linearly with a logarithmic decrease in the rate of deformation. Note, therefore, that the physical mechanism of this rate effect is completely different from that of the rate effect expressed by Eq. (1d). Note also that their effects on the shear traction are opposite, in the sense that the rate effect expressed by (1d) increases the shear traction with a decrease in the slip rate, whereas the rate effect observed for the shear fracture of intact rock decreases the shear traction with a decrease in the rate of deformation.

Although the rate effect observed for the shear fracture of intact rock has conventionally been expressed in terms of the strain rate $\dot{\varepsilon}$, it may equivalently be expressed in terms of $\dot{D}$ as:

$$\tau_p = g(\sigma_n^{\mathrm{eff}}) \left[1 + \gamma \ln \left(\frac{\dot{D}}{\dot{D} + \dot{D}_0} \right) \right] \quad (4)$$

because $\dot{\varepsilon}$ is directly proportional to $\dot{D}$. In the above equation, g is a function of σ_n^{eff}, $\dot{D}_0$ is the reference rate of slip (or relative displacement) along the shear fracture surfaces, and γ is a numerical constant. Equation (4) has been formulated so as for the rate effect to saturate when $\dot{D} \gg \dot{D}_0$. When $\dot{D} \ll \dot{D}_0$, equation (4) is reduced to:

$$\tau_p = g(\sigma_n^{\mathrm{eff}}) \left[1 + \gamma \ln \left(\frac{\dot{D}}{\dot{D}_0} \right) \right]. \quad (4')$$

The functional form $g(\sigma_n^{\mathrm{eff}})$ is expressed as a linear function of σ_n^{eff} as follows (see Ohnaka, 1995):

$$g(\sigma_n^{\mathrm{eff}}) = c_0 + c_1 \sigma_n^{\mathrm{eff}} \quad (5)$$

where c_0 and c_1 are constants. From laboratory experiments, we have $\gamma = 0.01$–0.02, $c_0 = 100$–140 MPa, and $c_1 = 0.7$–

0.75 for the shear fracture of intact granite (Masuda *et al.*, 1987; Kato *et al.*, 2003a, b; Ohnaka, 1995). We thus find that the rate effect expressed by Eq. (4) or (4') is also very small. Although the rate effect expressed by (4) or (4') has been found for intact rock, it should be noted that the same effect is possibly operative at asperity junctions of high stress concentration on a pre-existing fault in wet environments.

The other constitutive law parameters τ_i, $\Delta\tau_b$, D_a, and D_c may also depend on $\dot{D}$, σ_n^{eff}, T, and CE; however, their effects on τ_i, $\Delta\tau_b$, D_a, and D_c are not known at present, so that they will be left for a future study.

The shear rupture that proceeds on irregular rupturing surfaces is governed by not only nonlinear physics of the constitutive law but also geometric properties of the rupture-surface irregularity. Indeed, it has been found with laboratory experiments (Ohnaka, 2003) that the fundamental cause of the scaling property lies at the characteristic length λ_c, which is defined as the predominant wavelength that represents geometric irregularity of the rupturing surfaces. The constitutive law displacement parameters D_a and D_c scale with λ_c according to the following laws (Ohnaka, 2003):

$$D_a = c_2 D_c \tag{6}$$

$$D_c = \mathrm{K}(\Delta\tau_b/\tau_p)^m \lambda_c \tag{7}$$

where c_2, K , and m are dimensionless constants. Since D_a and D_c are the constitutive law parameters that scale with λ_c, the scaling property is automatically incorporated into the slip-dependent constitutive law. It has been shown that scaling of scale-dependent physical quantities inherent in the shear rupture is commonly reduced to the scale-dependence of D_c (Ohnaka and Shen, 1999; Ohnaka, 2000, 2003, 2004).

In general, a larger fault includes a geometrically larger asperity area of high resistance to rupture growth in a statistical sense, and the irregular rupture surfaces of such a geometrically larger asperity area contain a longer predominant wavelength component λ_c. Accordingly, λ_c is longer for a larger earthquake fault, and consequently, D_c is larger for a larger earthquake. For instance, such scale-dependent physical quantities as the length of the nucleation zone and its duration, the slip acceleration, and the apparent shear rupture energy G_c, scale with D_c, which in turn scales with λ_c according to Eq. (7) (Ohnaka and Shen, 1999; Ohnaka, 2003, 2004). Thus, λ_c plays a fundamental role in scaling scale-dependent physical quantities inherent in the shear rupture of a broad scale range (Ohnaka, 2003). It will be shown in Section 5 how the nucleation zone length and its duration scale with λ_c.

When the shear traction τ is specifically given as a function of D as shown in Fig. 1(a), the apparent shear rupture energy G_c is given by (Palmer and Rice, 1973):

$$G_c = \int_0^{D_c} [\tau(D) - \tau_r] dD. \tag{8}$$

As intuitively understood from the fact that the value of integral (8) is equal to the area of the hatched portion in Fig. 1(a), the slip-dependent constitutive law automatically satisfies the Griffith energy balance fracture criterion. In addition, as discussed above, the slip-dependent constitutive law can be formulated so as to incorporate both the rate property and the scaling property into itself. In this sense, the slip-dependent constitutive law is a more universal, physical law than the Griffith criterion. G_c represents the energy required for the rupture front to further grow, and hence the resistance to rupture growth is defined as G_c.

In this section, I have attempted a thorough discussion on what the constitutive law for earthquake ruptures ought to be from the standpoint of rock physics on the basis of solid facts observed in the laboratory. The arguments presented in this section lead to the conclusion that the governing law for earthquake ruptures should be formulated as a slip-dependent constitutive law with parameters that may be an implicit function of slip rate or time. In the subsequent sections, I will concentrate my attention on how the entire process of one cycle for a typical, large earthquake can be modeled, and then how consistently the process leading up to a large earthquake is modeled in the framework of fault mechanics based on the slip-dependent constitutive law, and the earthquake cycle model.

3. Cyclic Process for Typical Large Earthquakes

There is a pervasive hypothesis that the Earth's crust is in the state of perpetual self-organized criticality in which any small earthquake may cascade into a large event, and that earthquakes are in principle unpredictable catastrophes. The Gutenberg-Richter frequency-magnitude power law has been cited as evidence that the Earth's crust is in the self-organized critical state. If the Gutenberg-Richter frequency-magnitude power law were, in a strict sense, applicable to all earthquakes over the entire range of magnitude, it would mean that there is no characteristic length scale in the Earth's crust. There are, however, increasing amounts of evidence against this. It would suffice to give a few counterexamples, which will be shown below.

Global seismicity catalogs including large earthquakes indicate that large earthquakes do not fall on a Gutenberg-Richter frequency-magnitude power law curve (Engdahl and Villasenor, 2002). Paleoseismological data suggest that individual faults and fault segments tend to generate characteristic earthquakes having a relatively narrow range of magnitude, which do not follow the Gutenberg-Richter frequency-magnitude power law (Schwartz and Coppersmith, 1984; Sieh, 1996). These show that there is no doubt that a number of characteristic length scales exist in the Earth's crust, such as the depth of seismogenic layer, fault and/or its segment sizes. This is not in favor of the hypothesis that the Earth's crust is in the state of perpetual self-organized criticality.

An earthquake cannot occur anywhere in the Earth's crust, but can occur only in the region where an adequate amount of elastic strain energy as its driving force has been accumulated. Once an earthquake occurs in such a region, the elastic strain energy is necessarily consumed, and hence the stored energy in the region is lowered to a sub-critical level. When the size of earthquake is small, the energy released is restricted within a small region, so that the released energy may easily be restored to its critical level by such means as fault-fault interaction and/or dynamic stress transfer immediately after the event. When the size of earthquake is large, however, a large amount of the elastic strain energy stored in a wide region is consumed and is lowered to a sub-critical level. The large amount of energy released in a wide region cannot easily be restored to its critical level immediately after the event, even by means of fault-fault interaction and/or dynamic stress transfer. Tectonic loading due to perpetual slow plate motion is necessarily required for this. The next large earthquake therefore cannot occur in the region for a long time until an adequate amount of elastic strain energy is accumulated again up to its critical level with tectonic loading. In this respect, large earthquakes are distinctly different from small earthquakes.

Historic records and paleoseismological data suggest that large earthquakes, in particular those along plate boundaries, have occurred repeatedly, neither in clusters nor at random but quasi-periodically, on a single fault, and that average recurrence time intervals are well defined (Schwartz and Coppersmith, 1984; Sieh, 1996; Ishibashi and Satake, 1998; Utsu, 1998). For instance, large earthquakes occurring repeatedly on a quasi-periodic basis (1361, 1498, 1605, 1707, 1854, and 1946) along the plate interface are best documented in history in the Nankai region in the southwest of Japan. Whether temporal distribution of such large events on a specific fault is clustered or quasi-periodic can be checked statistically under the assumption that the sequence of identical events is independently distributed. If it is further assumed that the probability density function $w(\tau)$ of the time interval distribution is represented by the Weibull distribution: $w(\tau) = \alpha\beta\tau^{\beta-1}\exp(-\alpha\tau^{\beta})$ (α and β being constants), the probability $p(u|t)$ that the next event occurs in a time interval from t to $t + u$ is (Utsu, 1984, 1999, 2002)

$$p(u|t) = 1 - \exp\{-\alpha[(t+u)^{\beta} - t^{\beta}]\} \quad (9)$$

and the mean time interval $E[\tau]$ and its variance $V[\tau]$ are

$$E[\tau] = \alpha^{-1/\beta}\Gamma(1/\beta + 1) \quad (10)$$

$$V[\tau] = \alpha^{-2/\beta}\{\Gamma(2/\beta + 1) - [\Gamma(1/\beta + 1)]^2\}. \quad (11)$$

The time series of event occurrences on a single fault can be classified in terms of β into the following four cases (Utsu, 1998): $0 < \beta < 1$ if events are clustered, $\beta = 1$ if events are random, $\beta > 1$ if events are intermittent, and $\beta \to \infty$ if events are strictly periodic. Utsu (1998) showed in his simulation that events virtually occur periodically when $\beta > 10$, and quasi-periodically even when $\beta = 3 - 6$. For those large historical earthquakes along the plate interface in the Nankai region, the average recurrence interval with its standard deviation has been evaluated to be 117.1 ± 21.2 years, and $\beta = 6.09$ (Utsu, 1998). It can thus be concluded that the sequence of large historical earthquakes in the Nankai region have occurred repeatedly on a quasi-periodic basis.

The elastic strain energy builds up to the critical level much faster in the elastic medium along plate boundaries than in regions away from plate boundaries. Hence, it is expected that large earthquakes occur intermittently more often along plate boundaries than in regions away from plate boundaries, and that the recurrence time interval for interplate earthquakes is much shorter than that for intra-plate earthquakes in regions away from plate boundaries. The recurrence interval is on the order of 10^2 years for large earthquakes that occur along a plate boundary fault between two tectonic plates whose relative motion has a rate of a few cm/year. Such a typical example is the large historical earthquakes along the plate interface in the Nankai region mentioned earlier. In contrast, it has been inferred for intra-plate paleoearthquakes that the recurrence time interval is on the order of 10^3 to 10^4 years or longer, depending on the rate of tectonic strain buildup (e.g., Kumamoto, 1998; Matsuda, 1998). With slower loading rates, the length of the recurrence interval is affected more by factors other than the loading rate. Nevertheless, the observations indicate that the recurrence interval depends on the tectonic loading rate. This fact is incompatible with the hypothesis that the crust is in the perpetual critical state. If the crust were in the perpetual critical state, it should continue to have a potential to cause the next large earthquake even immediately after the occurrence of a large event. One might therefore expect that large earthquakes occur more often in the same region, irrespective of the tectonic loading rate. This, however, does not agree with the observations. Thus, the observations consistently lead to the conclusion that the hypothesis of perpetual self-organized criticality is at least not applicable to large earthquakes.

An alternative, more rational approach is to assume that the crust immediately after a large earthquake is in a subcritical state, and that crustal deformation proceeds toward the critical state with tectonic loading. An essential feature of this model is that the process from a sub-critical state to the critical state is repeated intermittently on a single fault under perpetual tectonic loading. We consider a system where an inhomogeneous, pre-existing fault that has a potential to cause a large earthquake is embedded in the brittle seismogenic layer, which is also inhomogeneous. Such a fault may be called a master fault. In particular, we specifically consider the deformation process of a crustal region including a master fault such as a plate boundary fault, from a subcritical state toward the critical state with tectonic loading, which can be regarded as the process leading up to a large earthquake.

The entire process of one cycle for a typical, large earthquake may commonly be modeled as shown in Fig. 2 (Ohnaka, 1998). Shortly after the occurrence of a large earthquake, the fault heals and is re-strengthened (phase I in Fig. 2), and hence the elastic strain energy can again be accumulated in the region surrounding the fault (phase II in Fig. 2), as tectonic stress builds up perpetually. The fault thus regains a potential to cause the next large earthquake.

At an early stage of phase II, the tectonic stress is far below the critical level, and the amount of the elastic strain energy accumulated in the region is inadequate. This stage is therefore characterized by quiescent seismicity (i.e., background seismicity), and it may be recognized as quiescent period of seismicity when attention is paid to its time domain, and as seismic gap when attention is paid to its spatial domain. Indeed, the well-known concept of seismic gaps (Imamura, 1928/29; Fedotov, 1965; Mogi, 1968; Sykes and Nishenko, 1984) has been proposed from seismicity studies. As the tectonic stress reaches higher levels, and gradually approaches its critical level, the crust begins to behave in-elastically, and consequently, seismicity in the region surrounding the master fault of the next large event becomes progressively ac-

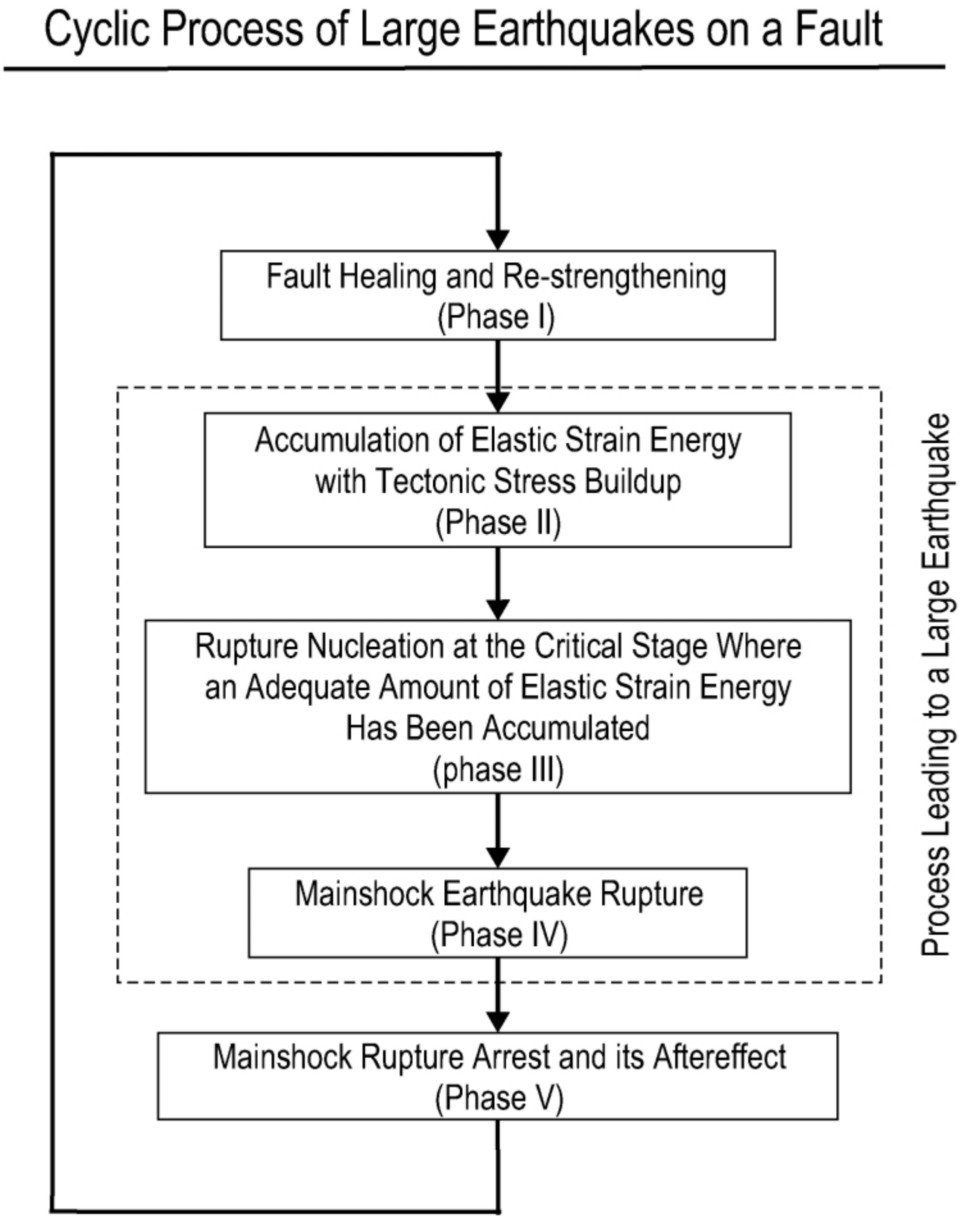

Fig. 2. A model of the cyclic process for large earthquakes.

tive with time, as shown schematically in Fig. 3. This is because the crust is inherently inhomogeneous and includes numerous faults of small-to-moderate sizes. This later stage of phase II can thus be characterized by activation of seismicity.

Seismic activity is enhanced by fluid-rock interaction, and triggering effect by stress transfer at the later stage of phase II where the tectonic stress is in close proximity to the critical state. For instance, a small stress perturbation due to the occurrence of neighboring earthquakes and/or the Earth tides may trigger small to moderate earthquakes at this stage. Indeed, a recent, elaborate study (Tanaka *et al.*, 2002) strongly suggests that a small stress change due to the Earth tide can trigger earthquakes in the region where the tectonic stress is in close proximity to the critical state. Activated seismicity at this stage may be recognized as premonitory phenomena for the ensuing mainshock earthquake.

Eventually, rupture nucleation begins to proceed locally at a place where the resistance to rupture growth is the weakest on the master fault (phase III in Fig. 2), when the tectonic stress has reached its critical level, and when an adequate amount of the elastic strain energy has been accumulated. The nucleation necessarily leads to the ensuing mainshock rupture on the fault (phase IV in Fig. 2), accompanied by rapid stress drop and dissipation of a great amount of the elastic strain energy, resulting in radiation of seismic waves. The arrest of the mainshock rupture results in its aftereffect that includes re-distribution of local stresses on and around the fault, leading to aftershock activity (phase V in Fig. 2).

Using the JMA earthquake catalogue data from 1977 through 1997, Maeda (1999) carefully examined temporal and spatial distributions of representative foreshocks as functions of the time from the mainshock origin time and the distance from the mainshock location, and he found that immediate (roughly within one day) foreshocks concentrate in the vicinity of the hypocenter of the pending mainshock

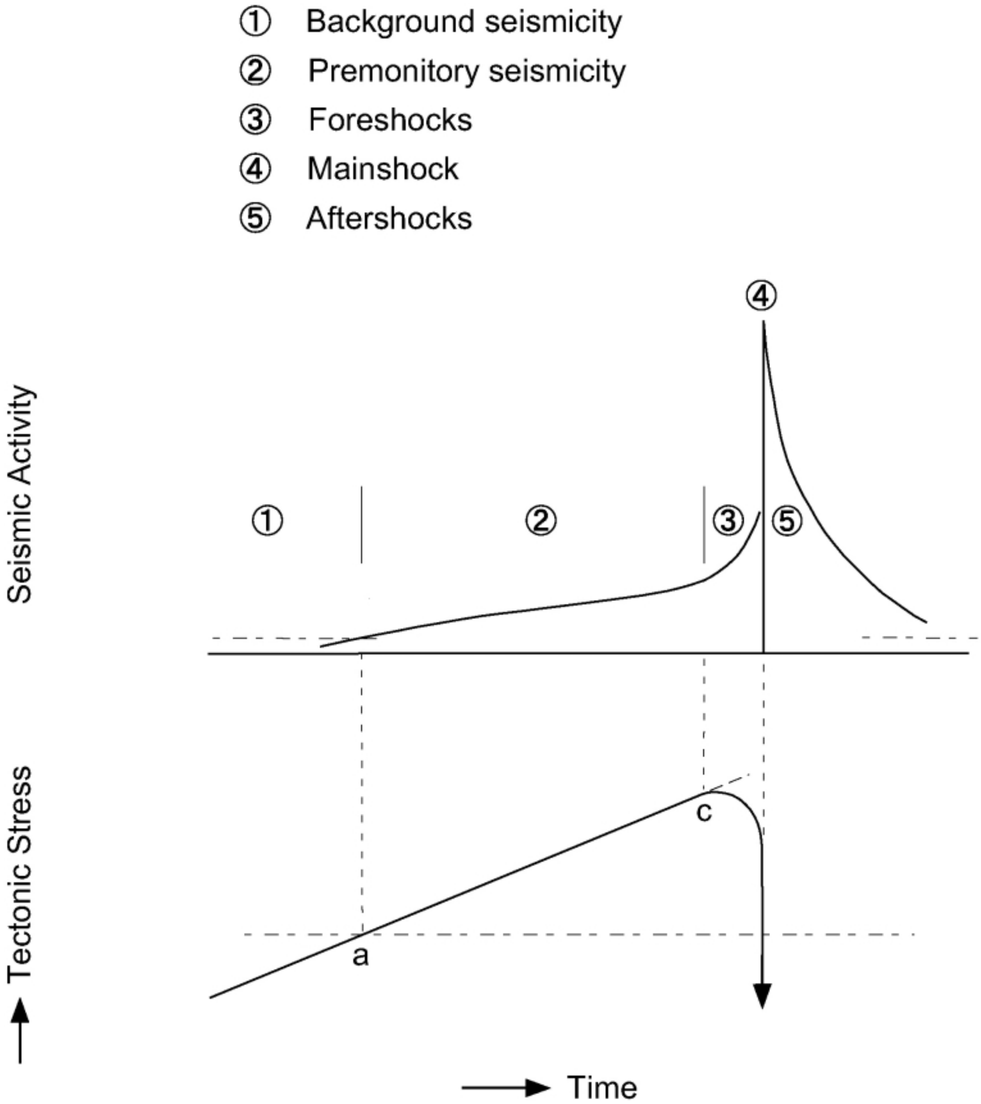

Fig. 3. A model of seismic activity enhanced during the deformation process leading up to a large earthquake. Seismic activity occurs after tectonic stress exceeds a threshold level (point a in the figure), and thereafter seismicity becomes progressively active with time. Mainshock rupture nucleation begins to proceed at point c in the figure.

earthquake. This prominent feature is well explained by a model of the rupture nucleation (Ohnaka, 1992) put forward based on laboratory experiments (Ohnaka *et al.*, 1993), which demonstrates that micro-seismic activities are indeed induced during the slip failure nucleation that proceeds on an inhomogeneous fault. During the nucleation of a typical, large earthquake, a region of a few kilometers in dimension (referred to as the nucleation zone) on the seismogenic fault would slip, initially at a very slow and steady rate and subsequently at accelerating rates, over a distance of the order of 1 m. Such a slip at an initially slow and steady rate, and subsequently at accelerating rates over a distance of the order of 1 m proceeds with or without inducing micro-earthquakes, depending on the fault structure and such ambient conditions as temperature (Ohnaka, 2000). If, for instance, a non-uniform fault in the brittle regime has sizable asperity patches on which irregularities (or micro-asperities) of short wavelength components are superimposed, slow slip failure in an asperity patch will necessarily bring about fracture of micro-asperities in the patch. In this case, the nucleation process carries micro-earthquakes (i.e., immediate foreshocks) (Fig. 3). Note that fracture of micro-asperities during the nucleation occurs despite the fact that the overall shear stress decreases (because of slip-weakening) in the nucleation zone (Ohnaka *et al.*, 1993). Immediate foreshock activities induced during the mainshock nucleation have been discussed for such earthquakes as the 1978 Izu Oshima Kinkai earthquake (Ohnaka, 1992, 1993), the 1983 Central Japan Sea earthquake (Shibazaki and Matsu'ura, 1995), and the 1992 Landers earthquake (Dodge *et al.*, 1995, 1996).

The above is a model of the cycle process for a typical, large earthquake that takes place in the brittle seismogenic layer characterized by heterogeneities. In this model, phases I to V come in cycles. In particular, phases II and III are integral parts of the process leading up to a typical, large earthquake that inevitably occurs in the brittle seismogenic layer. Since the present paper is concerned with a physical model with predictive capability for a typical, large earthquake, we need to focus on phases II and III of the cycle process. It has been argued by earlier authors (Imamura, 1928/29; Fe-

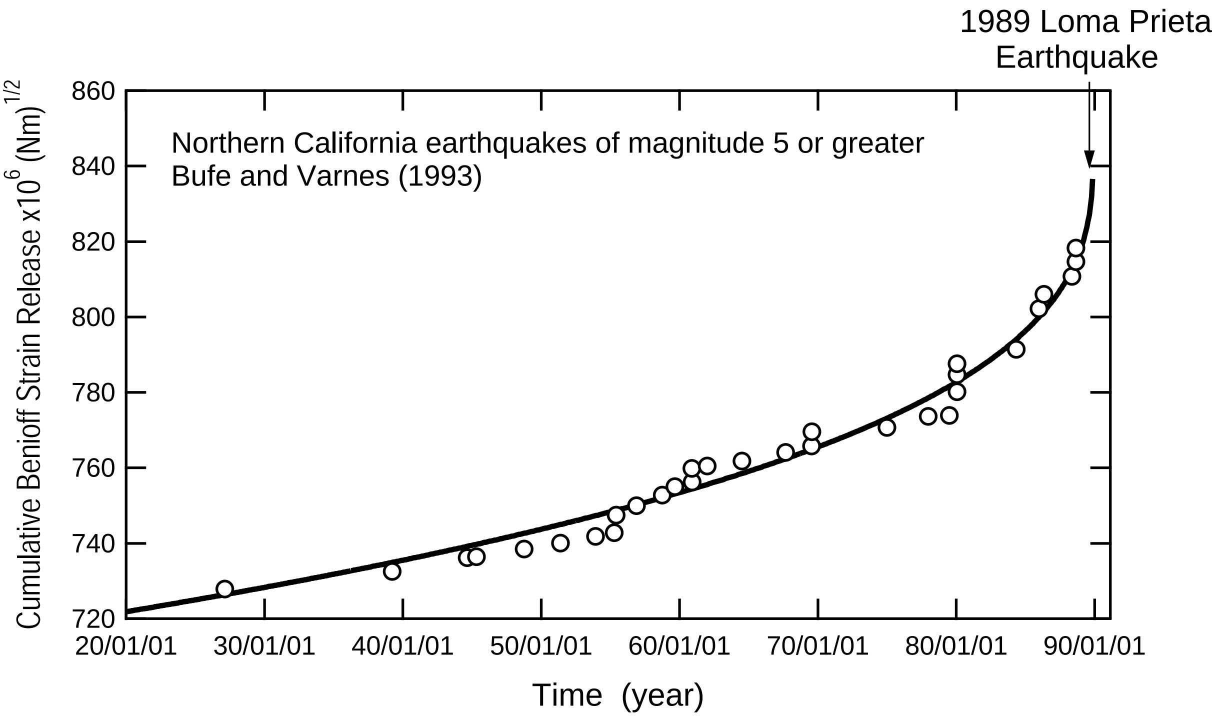

Fig. 4. A plot of the cumulative Benioff strain release against time. Open circles indicate data points of earthquakes with magnitude 5 or greater, and a thick curve represents the result of the best-fit regression analysis. Reproduced from a paper by Bufe and Varnes (1993).

dotov, 1965; Mogi, 1968; McCann *et al*., 1979; Sykes and Nishenko, 1984; Nishenko, 1991) that seismic gaps, which are interpreted as quiescent seismicity at an early stage of phase II in the context of the present model, are useful for long-term forecasting. For more information about long-term forecasts based on the concept of seismic gaps, see a recent paper by Kanamori (2002), who provides an overview of the status quo and problems of earthquake prediction in general. In the subsequent sections, our focus will be directed on the later stage of phase II for physical modeling of intermediate-term forecasting, and on phase III for physical modeling of short-term (or immediate) forecasting.

4. Physical Modeling for Intermediate-Term Forecasting

I here concentrate on how the process leading up to a large earthquake in phase II can be modeled in the context of the seismic cycle model presented above. Phase II defined above can be regarded as the preparation process for a large earthquake, in the sense that the elastic strain energy as the driving force to bring about a next large earthquake, builds up in this phase with tectonic loading, and also in the sense that the deformation of the crust proceeds toward the catastrophic failure with tectonic loading. In view of this property of phase II, models for intermediate-term forecasting have been proposed by earlier authors.

For instance, the well-known time-to-failure function model (e.g., Bufe and Varnes, 1993; Bufe *et al*., 1994; Jaume and Sykes, 1999) is a typical model for intermediate-term forecasting. The model focuses on accelerating seismic activity observed during the deformation process of the inhomogeneous crust leading up to a major earthquake with tectonic loading in phase II, and is based on the laboratory observation that the deformation process of an inhomogeneous body leading to the ensuing catastrophic failure obeys a power law of the form (Saito, 1969; Varnes, 1989):

$$\frac{d\Omega(t)}{dt} = \frac{k}{(t_f - t)^n} \tag{12}$$

where $\Omega(t)$ represents strain or any of measurable quantities (such as event count, Benioff strain, or seismic moment) as a function of time t, t_f represents the time of failure, and k and n are constants. Integration of (12) leads to:

$$\Omega(t) = \Omega_f - \frac{k}{1-n}(t_f - t)^{1-n} \qquad (n \neq 1) \tag{13}$$

where $\Omega_f = \Omega(t_f)$.

Figure 4, taken from a paper by Bufe and Varnes (1993), shows a plot of $\Omega(t)$ against time t for the period 1855–1989 for northern California earthquakes of magnitude 5 or greater. $\Omega(t)$ in this figure is defined as the cumulative Benioff strain release, which is given by (Bufe and Varnes, 1993):

$$\Omega(t) = \sum_{i=1}^{N(t)} \omega_i \tag{14}$$

where ω_i is the Benioff strain release for each event, calculated from $\log \omega_i = 0.75M_i + d$. Here, M_i represents the magnitude of the i-th event in the sequence of N earthquakes that occurred in a certain region, and d is a constant. The Loma Prieta earthquake (M7.1) occurred on October 18, 1989, and one can see from Fig. 4 that seismic activity in the region accelerated progressively towards the 1989 Loma Prieta event, and that the accelerating seismicity is well explained by the power law of Eq. (13).

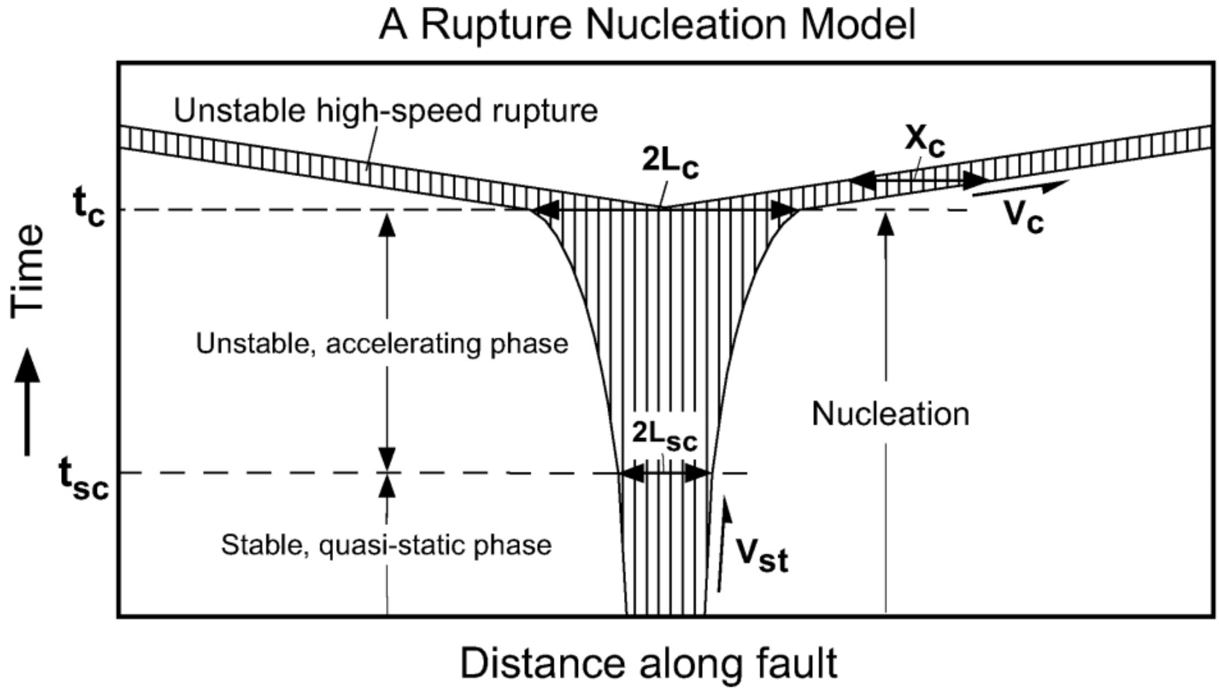

Fig. 5. A model of rupture nucleation. The rupture begins to grow stably at a steady, slow speed V_{st} to a critical length $2L_{sc}$ (at $t = t_{sc}$), from which it extends spontaneously at accelerating speeds up to another critical length $2L_c$ (at $t = t_c$). Beyond the critical length $2L_c$, the rupture propagates at a steady, high speed V_c close to the shear wave velocity. The hatched portion represents the zone in which the breakdown (or slip-weakening) proceeds with time. X_c denotes the breakdown zone length, and $2L_c$ denotes the critical length of the nucleation zone.

Both t_f and n in Eq. (13) can be evaluated from the best-fit regression analysis using earthquake events that have occurred prior to the imminent mainshock to be predicted, and the magnitude of the expected event at $t = t_f$ can also be estimated when $n < 1$ (Bufe and Varnes, 1993). The thick curve in Fig. 4 represents the result of the best-fit regression analysis made using earthquakes of magnitude 5 or greater that occurred during the period 1927–1988 (Bufe and Varnes, 1993). One can see from Fig. 4 that the time-to-failure function model is useful for intermediate-term forecasting. The model seems to well explain accelerating seismic activity observed during the process leading up to a major event in other regions as well (Bufe *et al.*, 1994; Brehm and Braile, 1998; Jaume and Sykes, 1999; Yin *et al.*, 2000).

Another model which may also be useful for intermediate-term forecasting is the load-unload response ratio (LURR) model (e.g., Yin *et al.*, 1995, 2000). At an early stage of phase II, the crust behaves elastically. However, it progressively behaves in-elastically as the tectonic stress approaches its critical level. If, therefore, a measure of the proximity to the critical state in a region is suitably defined, the imminent earthquake rupture in the region may be predicted. LURR is practically defined as the ratio of the cumulative Benioff strain release during loading to that during unloading as determined by calculating Earth tide induced perturbations in the Coulomb failure stress on optimally oriented faults (Yin *et al.*, 2000). LURR thus defined represents a measure of the proximity to the critical state, and high LURR values (> 1) indicate that the region is in close proximity to the critical state that has a potential to cause a major earthquake. Values for LURR have been calculated for various regions of different tectonic regimes (Yin *et al.*, 2000, 2002), and the results indicate that the model provides a useful means for intermediate-term forecasting (Yin *et al.*, 2000, 2002).

Both models of the time-to-failure function and the load-unload response ratio are based on the common fact that inelastic deformation of the crust develops with tectonic loading in phase II prior to the occurrence of a major earthquake. It is therefore not by coincidence that the scaling relation between the critical region size and the magnitude of the final event (Bowman *et al.*, 1998), estimated from the time-to-failure function model, agrees with that estimated from the load-unload response ratio model (Yin *et al.*, 2002). The empirical scaling relation found by Bowman *et al.* (1998) can be approximated in terms of the critical region radius R_c and the rupture area S of the final event as follows (Rundle *et al.*, 2000):

$$R_c = 10S^{1/2}. \qquad (15)$$

This relation indicates that the preparation process for a larger earthquake develops in a wider region, and that accelerating seismic activity during the deformation process of the crust leading up to a major event occurs in a wide region with the critical radius R_c being ten times larger than the characteristic fault length $S^{1/2}$.

5. Physical Modeling for Short-Term (Immediate) Forecasting

The power law of Eq. (12) or (13) proposed for intermediate-term forecasting possibly has a singularity at $t = t_f$, depending on a value evaluated for the exponent n. The possible singularity at $t = t_f$ comes from the fact that equation (12) has been derived without considering the physical process immediately before the imminent event we

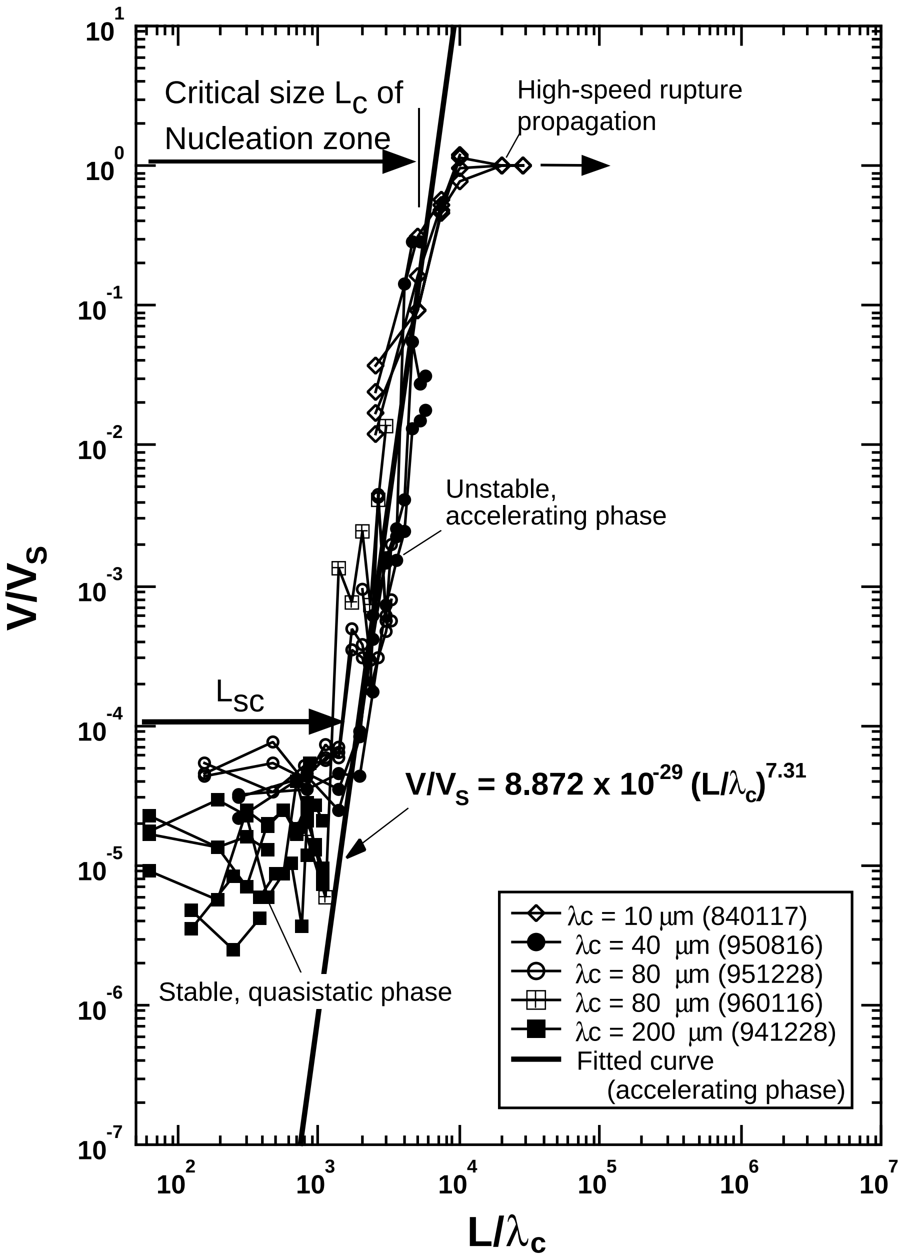

Fig. 6. A plot of the logarithm of the rupture growth rate V normalized to the shear wave velocity V_S against the logarithm of the rupture growth length L normalized to the critical length λ_c for shear failure nucleation on faults with different surface roughnesses. Reproduced from a paper by Ohnaka and Shen (1999).

intend to predict. Eq. (12) or (13) may therefore not be suitable for short-term (immediate) forecasting, though it is certainly useful for intermediate-term forecasting.

If we are concerned with physical modeling of the process leading up to a large event for short-term (or immediate) forecasting, we have to direct our focus on the physical process or phase immediately before the imminent event we intend to predict. The preparation process immediately before an imminent earthquake rupture is what is referred to as the nucleation process (phase IV in Fig. 2). It is therefore crucial to incorporate the nucleation process into a physical model for short-term (or immediate) forecasting.

An earthquake rupture (or unstable, dynamic shear rupture) on a fault characterized by inhomogeneities cannot begin to propagate abruptly at speeds close to sonic velocities from a stable and static state, but is necessarily preceded by a stable and quasi-static phase of rupture nucleation and the subsequent accelerating phase. A physical model of such nucleation that proceeds on an inhomogeneous fault has been proposed for typical earthquakes (Fig. 5), based on the results revealed in the high-resolution laboratory experiments (Ohnaka *et al.*, 1986; Ohnaka and Kuwahara, 1990; Ohnaka and Shen, 1999; Ohnaka, 2000, 2004). In the framework of fault mechanics based on the slip-dependent constitutive law, theoretical studies on rupture nucleation and its numerical simulations have also been done (Yamashita and Ohnaka, 1991; Matsu'ura *et al.*, 1992; Shibazaki and Matsu'ura, 1992, 1995, 1998; Ionescu and Campillo, 1999; Bizzarri *et al.*, 2001; Uenishi and Rice, 2003). In particular, Shibazaki and Matsu'ura (1998) clearly showed in their

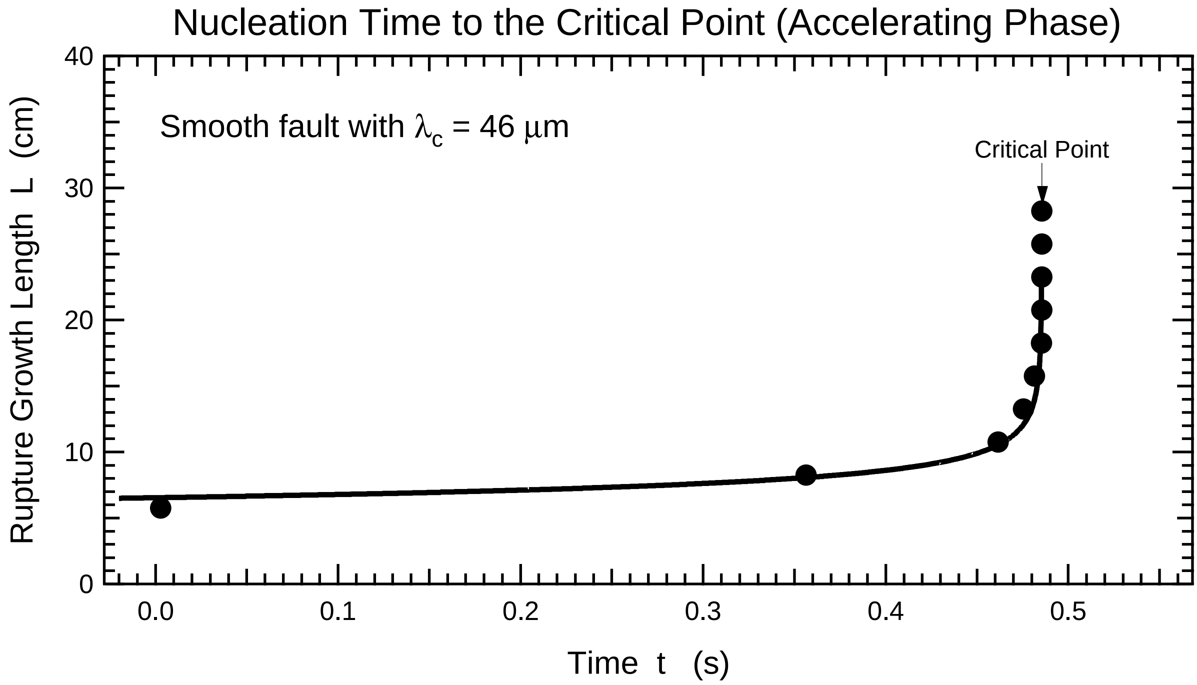

Fig. 7. A plot of the rupture growth length $L(t)$ against time t for the accelerating phase of nucleation that occurred on a fault with $\lambda_c = 46$ μm. Black circles denote data points, and a thick curve denote Eq. (17) fitted to data points. In the figure, the origin of time t is taken such that $t_{sc} = 0$. Slightly modified from figure 3 in a paper by Ohnaka (2004).

numerical simulations that the nucleation process observed in laboratory experiments can be reproduced completely, if such non-uniform distributions of the constitutive law parameters D_c and $\Delta\tau_b$ as determined from the laboratory experiments are given specifically along a simulated fault. This corroborates the findings in laboratory experiments on rupture nucleation.

The shear rupture nucleates at a place where the resistance to rupture growth is the weakest, and it proceeds stably at a steady, slow speed V_{st} to a critical length L_{sc} (half-length), beyond which the rupture grows spontaneously at accelerating speeds to another critical length L_c (half-length), obeying a power law of the form (Fig. 6):

$$V/V_S = \alpha(L/\lambda_c)^n \tag{16}$$

where V is the rupture growth velocity, V_S is the shear wave velocity, L is the rupture growth length, λ_c is the characteristic length defined as the predominant wavelength that represents geometric irregularity (or roughness) of the rupturing surfaces in the slip direction, and α and n are numerical constants ($\alpha = 8.87 \times 10^{-29}$, and $n = 7.31$; see Ohnaka and Shen, 1999). Beyond the critical length L_c, the rupture propagates at a steady, high speed V_c close to elastic wave velocities (Fig. 5).

The nucleation process for $L < L_{sc}$ is a quasi-static rupture growth controlled by the rate of an applied load, whereas the process for $L_{sc} < L < L_c$ is the subsequent, spontaneous rupture growth driven by the release of the elastic strain energy stored in the surrounding medium (Ohnaka and Shen, 1999). The behavior of rupture growth therefore changes at $L = L_{sc}$ from a quasi-static phase controlled by the loading rate to a self-driven, accelerating phase controlled by the inertia. The behavior of the rupture also changes at $L = L_c$ from the accelerating phase to the phase of a steady propagation at a high-speed V_c. We focus on the critical length L_c, because the nucleation zone size estimated from seismological data corresponds to L_c (Ellsworth and Beroza, 1995; Shibazaki and Matsu'ura, 1998). Note therefore that the critical length L_c to be discussed below is physically not identical with the critical length defined by Andrews (1976a, b).

From Eq. (16), we derive a law that governs the nucleation process leading to the critical point. The critical point is defined specifically as the critical time t_c at which $L = L_c$ is attained. Noting that $dt = dL/V$, we have from (16) the following relation (Ohnaka, 2004):

$$L(t) = L_c \left(\frac{t_a - t_c}{t_a - t}\right)^{1/(n-1)} \qquad (t \le t_c) \tag{17}$$

where the origin of time t is taken such that the time t_{sc} (at which $L = L_{sc}$) = 0, and t_a is defined by:

$$t_a = \frac{1}{\alpha(n-1)}\frac{\lambda_c}{V_S}\left(\frac{\lambda_c}{L_c}\right)^{n-1} + t_c. \tag{18}$$

As stated previously, real rupture surfaces of heterogeneous materials such as rock cannot be flat planes, but they inherently exhibit geometric irregularity. Although these irregular rupture surfaces exhibit self-similarity, they cannot be self-similar at all scales but are self-similar within a finite scale range. This is because the shear rupture process is the process that smoothes away geometric irregularities of the rupturing surfaces. In general, shear rupture surface topographies that exhibit band-limited self-similarity can be quantified and characterized by two quantities: fractal dimension and corner wavelength. The corner wavelength is defined as the critical wavelength that separates the neighboring two

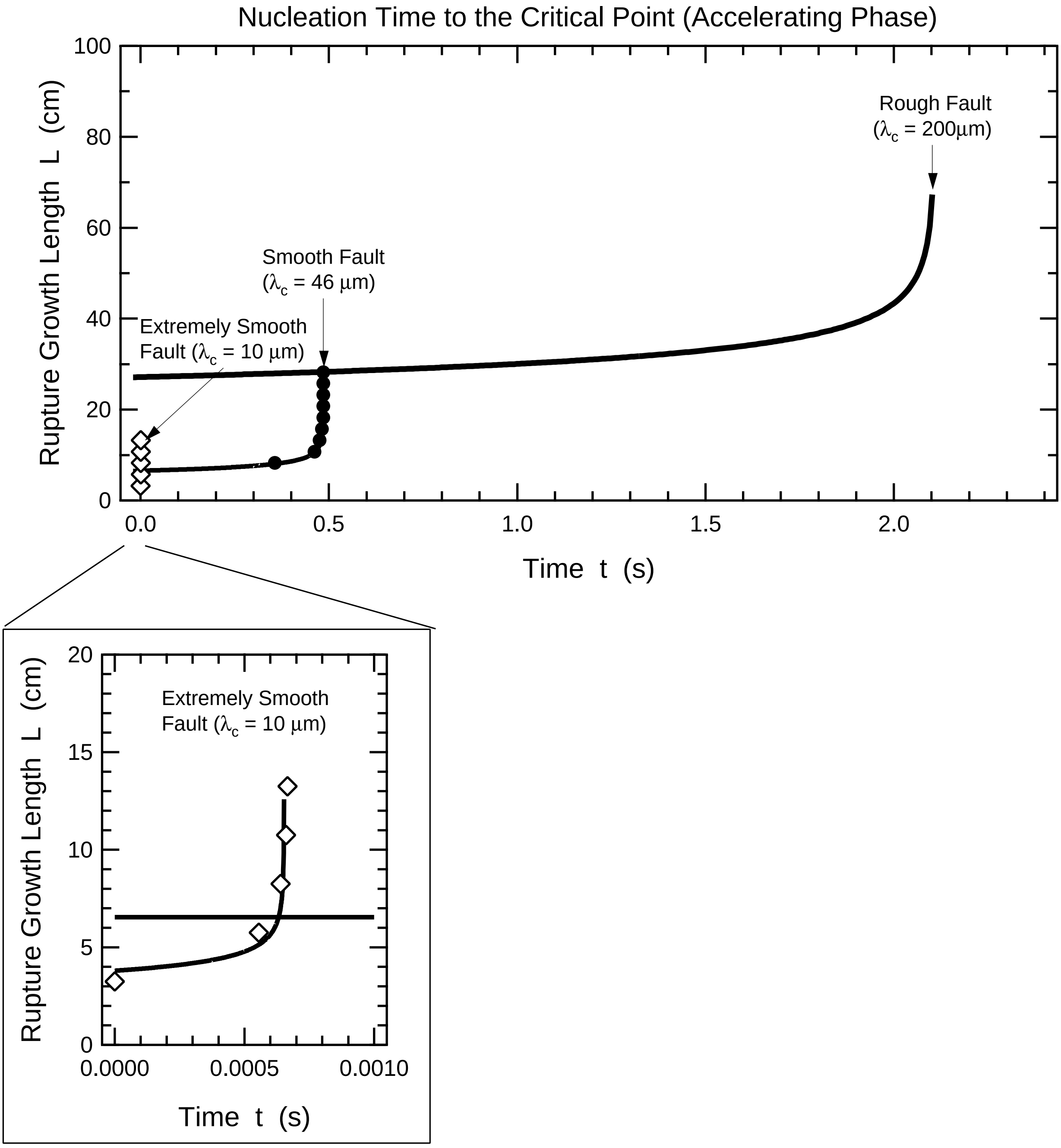

Fig. 8. A comparison of the relation between the rupture growth length $L(t)$ and time t for the accelerating phase of nucleation that occurred on faults with different surface roughnesses. The relation between $L(t)$ and t for nucleation on an extremely smooth fault with $\lambda_c = 10\ \mu$m is enlarged in the inset. In the figure, the origin of time t is taken such that $t_{sc} = 0$.

bands with different fractal dimensions. Of these two quantities, only the corner wavelength represents a characteristic length λ_c in the slip direction on the fault. The characteristic length determined from the corner wavelength is the predominant wavelength representing geometric irregularity (roughness) of the rupture surfaces in the slip direction (see Ohnaka and Shen, 1999; Ohnaka, 2003).

Figure 7 exemplifies how well equation (17) explains laboratory data on frictional slip failure nucleation in quantitative terms (Ohnaka, 2004). In this figure, the rupture growth length L is plotted against time t for data on the nucleation tested on a pre-cut fault with the surface roughness characterized by $\lambda_c = 46\ \mu$m. One can clearly see from Fig. 7 that equation (17) explains well laboratory data on the nucleation in quantitative terms. This corroborates the theoretical result that the nucleation process leading up to the critical point obeys the power law of Eq. (17).

In Fig. 8, data on the nucleation process leading up to the critical point tested on pre-cut faults with three different surface geometric irregularities (or roughnesses) are compared. The fault surface roughnesses used are reproduced in Fig. 9 for comparison: the rough surface characterized by the predominant wavelength of $\lambda_c = 200\ \mu$m, the smooth surface characterized by $\lambda_c = 46\ \mu$m, and the extremely smooth surface characterized by $\lambda_c = 10\ \mu$m (see Ohnaka and Shen, 1999).

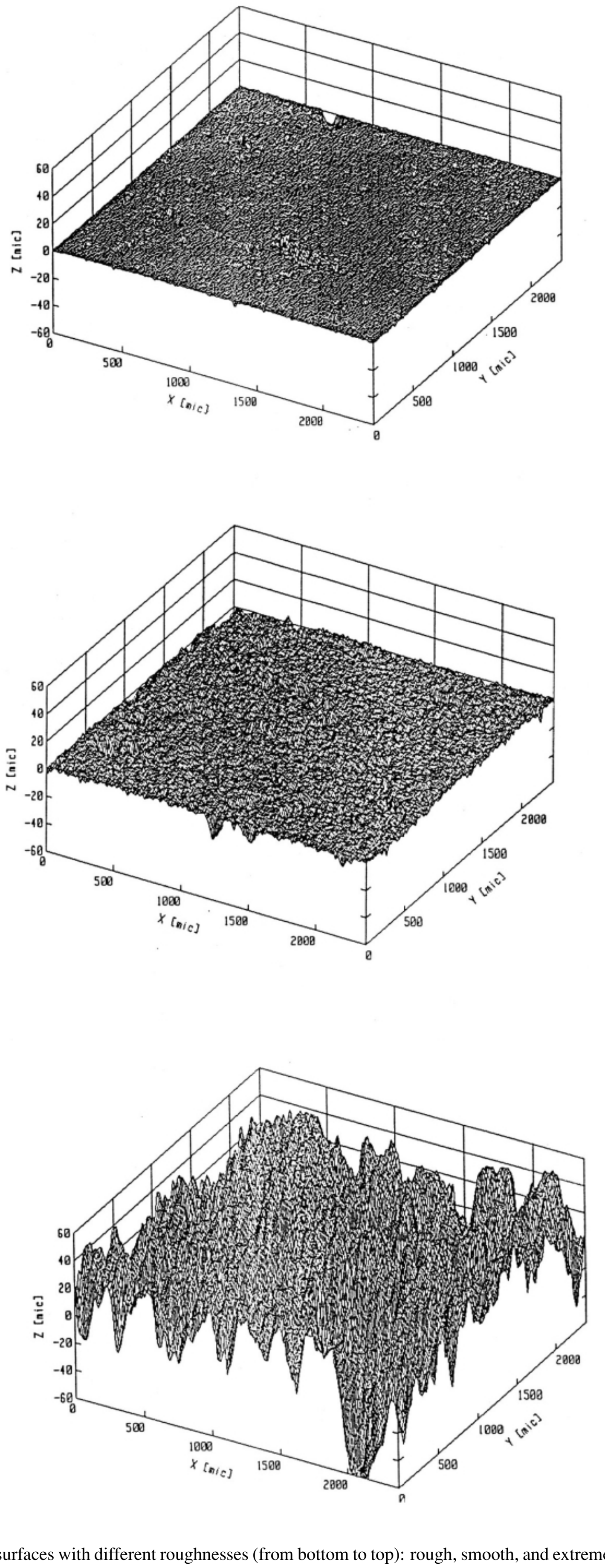

Fig. 9. A comparison of three fault surfaces with different roughnesses (from bottom to top): rough, smooth, and extremely smooth. The rough surface was prepared with grit #60, the smooth surface was prepared with grit #600, and the extremely smooth surface was prepared with grit #2000. Reproduced from a paper by Ohnaka and Shen (1999).

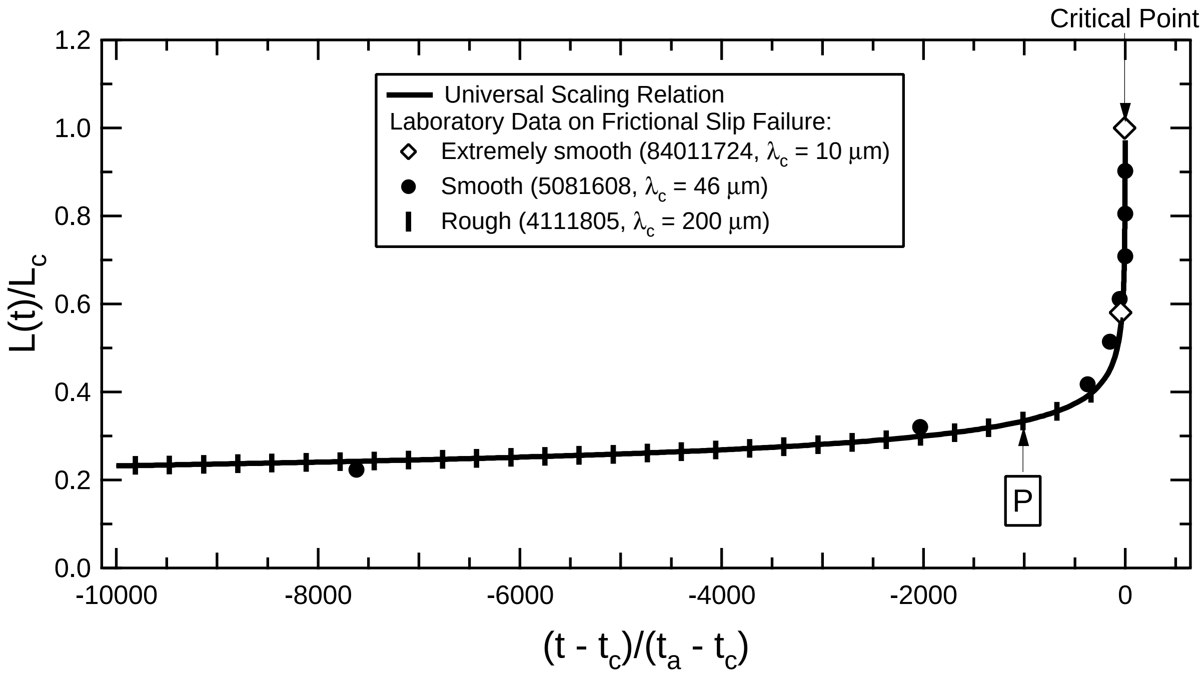

Fig. 10. A plot of $L(t)/L_c$ against $(t - t_c)/(t_a - t_c)$. A thick curve denotes Eq. (19). Laboratory data tested on faults with different surface roughnesses are also plotted for comparison.

It is clear from Fig. 8 that the rupture growth length and its duration (time to the critical point) are very long for the nucleation that proceeds on the rough fault with $\lambda_c = 200$ μm, if compared with those for the nucleation that proceeds on a very smooth fault with $\lambda_c = 10$ μm (inset). It can thus be concluded from Fig. 8 that both the rupture growth length L and the duration $|t - t_c|$ during the nucleation depend on the characteristic wavelength (or predominant wavelength) λ_c representing geometric irregularity of the fault surfaces, and that both L and $|t - t_c|$ increase systematically with an increase in λ_c. This indicates that the rupture growth length and its duration are scale-dependent, and that λ_c plays a crucial role in scaling the nucleation process.

The power law of Eq. (17) explains well experimental data (Fig. 7); however, this mathematical expression is scale-dependent (Fig. 8). To derive a scale-independent, universal expression, we rewrite Eq. (17) as follows (Ohnaka, 2004):

$$\frac{L(t)}{L_c} = \left(\frac{1}{1 - (t - t_c)/(t_a - t_c)} \right)^{1/(n-1)} \tag{19}$$

where

$$\frac{t - t_c}{t_a - t_c} = \alpha(n - 1) \left(\frac{L_c}{\lambda_c} \right)^{n-1} \frac{t - t_c}{(\lambda_c / V_S)}. \tag{20}$$

Expression (19) has a mathematical form that the relation between $L(t)/L_c$ and $(t - t_c)/(t_a - t_c)$ is scale-invariant, if n is scale-invariant. In fact, experimental data on frictional slip failure nucleation shown in Fig. 8 can be unified completely in quantitative terms by this expression (Fig. 10).

Figure 10 shows a plot of $L(t)/L_c$ against $(t - t_c)/(t_a - t_c)$ for the data on frictional slip failure nucleation shown in Fig. 8. Theoretical relation (19) is also over-plotted in Fig. 10 for comparison with the experimental data. One can see from Fig. 10 that different sets of experimental data tested on faults with different λ_c are unified completely in quantitative terms by expression (19). Since $t - t_c$ scales with λ_c / V_S (see Eq. (20)), it is obvious from Fig. 10 that not only the rupture growth length but also the nucleation time to the critical point scale with the characteristic length λ_c. We can thus confirm that the characteristic length λ_c plays a fundamental role in scaling not only the nucleation zone length but also the nucleation time to the critical point.

6. Physical Scaling of the Nucleation Process from Laboratory-Scale to Field-Scale

I have shown in previous section that both the nucleation zone length and the nucleation time to the critical point scale with the characteristic length λ_c that represents geometric irregularity of the rupturing surfaces. This poses questions about how long the effective characteristic length λ_c, the critical length L_c of the nucleation zone, and the nucleation time t_c to the critical point are for real, typical large earthquakes.

The fact that the rupture growth length L scales with the characteristic length λ_c implies that the size of mainshock earthquake scales with its nucleation zone size. Indeed, a physical scaling relation between mainshock seismic moment M_0 and its nucleation zone length $2L_c$ (L_c, critical half length) can be derived theoretically from a laboratory-based slip-dependent constitutive law as follows (Ohnaka, 2000):

$$M_0 = c_1 c_2 (k\kappa\Gamma/4)^3 (S_{A1}/aS)^3 (\Delta\tau_b/\overline{\Delta\tau})^6 \mu (2L_c)^3 \tag{21}$$

under the following assumptions (Ohnaka, 2000): 1) patches of high resistance to rupture growth on an inhomogeneous fault are so tough that an adequate amount of the elastic strain energy is accumulated in the medium surrounding the patches, 2) the rest of the fault is so weak that little amount

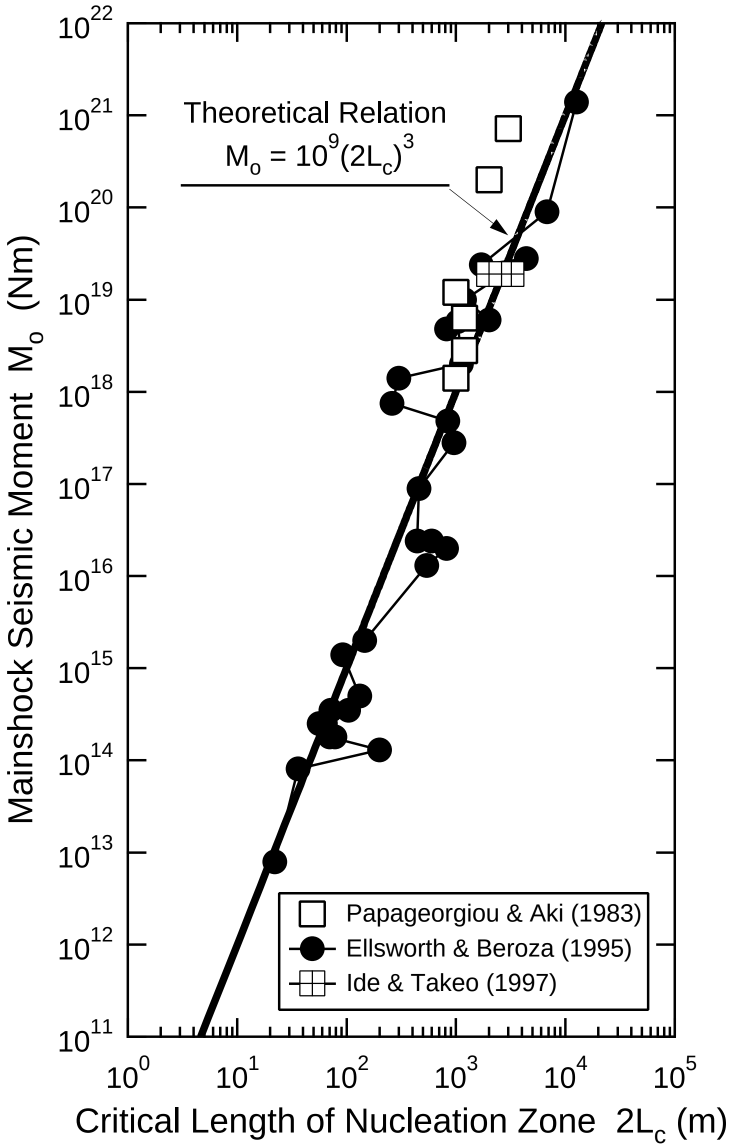

Fig. 11. A plot of the logarithm of the seismic moment M_0 against the logarithm of the critical length $2L_c$ of the nucleation zone for earthquakes. The theoretical scaling relation denoted by a thick line is compared with seismological data. Reproduced from figure 6 in a paper by Ohnaka (2000).

of the elastic strain energy is accumulated, and 3) the interaction between patches of high resistance to rupture growth is negligible. In Eq. (21), S_{A1} is the area of the geometrically largest patch of high resistance to rupture growth, S is the mainshock fault area, $\Delta\tau_b$ is the breakdown stress drop, $\overline{\Delta\tau}$ is the stress drop averaged over S, μ is the rigidity, and c_1, c_2, k, κ, Γ, and a are dimensionless constants (see Ohnaka, 2000).

Under appropriate assumptions, equation (21) is reduced to (Ohnaka, 2000):

$$M_0 = 1.0 \times 10^9 (2L_c)^3. \quad (22)$$

This theoretical scaling relation shows that the ensuing mainshock seismic moment is proportional to the 3rd power of the critical length of its nucleation zone. Equation (22) explains well data (Ellsworth and Beroza, 1995; Ohnaka, 2000) on earthquake nucleation in quantitative terms (Fig. 11). In the nucleation model shown in Fig. 5, the nucleation zone length L_c equals the breakdown zone length X_c, and contemplating that L_c is of the order of X_c, the data on X_c and M_0 for earthquakes analyzed by Papageorgiou and Aki (1983) are also plotted in Fig. 11 for comparison.

Figure 12 schematically shows an asperity model used for deriving Eq. (21). The term "asperity" is defined in this paper as a local area (or patch) of high resistance to rupture growth on a fault. A stable and slow growth of rupture, which may initially have been caused by tectonic loading, cannot begin to propagate spontaneously and dynamically, unless at least one of the patches of high resistance to rupture growth is broken down by the stable and slow growth of rupture. This is because there is no amount of energy available as a driving force to bring about dynamic rupture, unless the stored elastic strain energy is released by the breakdown of any patch of high resistance to rupture growth. In other words, the breakdown of one of those patches due to a stable and slow rupture growth is the prerequisite for dynamic high-speed rupture. Such an initial, stable and slow growth of patch rupture is what is called the nucleation process.

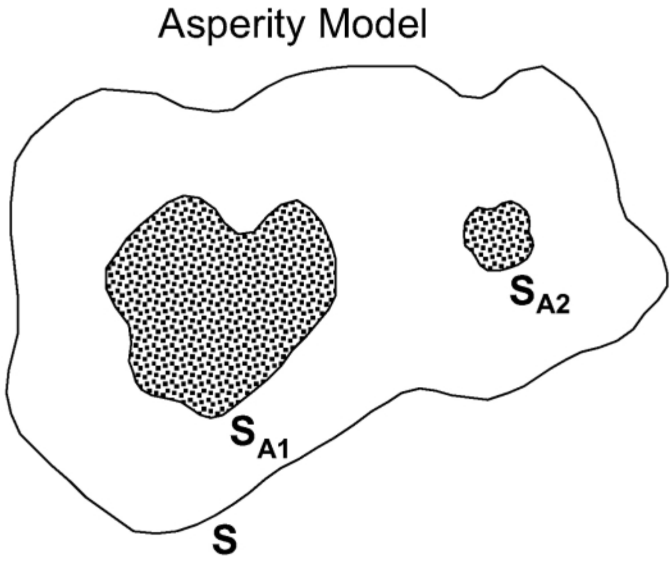

Fig. 12. Schematic diagram of an asperity model. S denotes fault area, and S_{A1} and S_{A2} denote the areas of patches of high resistance to rupture growth on the fault.

If the size of an initial, broken-down patch is geometrically large, the amount of the elastic strain energy to be released is large, so that the ensuing earthquake will be large. On the other hand, a large amount of D_c is by definition required for the breakdown of a geometrically large patch. The critical length L_c of nucleation zone is related to D_c by (Ohnaka, 2000)

$$L_c = \frac{1}{k}\frac{\mu}{\Delta\tau_b}D_c \quad (23)$$

where k is a dimensionless parameter. This leads to the conclusion that a large amount of D_c necessarily results in a large size of the nucleation zone. The above explanation provides rational, physical grounds for scaling law (21) or (22).

If, however, the interaction between patches is not negligible, scaling law (21) or (22) may no longer hold. Consider a case where the amount of the elastic strain energy released by the breakdown of a small patch (for instance, A2 in Fig. 12) is adequate enough to break down a neighboring large patch (A1 in Fig. 12). In this case, the size of the ensuing mainshock earthquake will be determined by the amount of the elastic strain energy released by the breakdown of this large patch (A1 in Fig. 12). On the other hand, the nucleation zone size is prescribed by the size of the small patch (A2 in Fig. 12) that has initially been broken down. Thus, the eventual size of mainshock earthquake may not necessarily scale with its nucleation zone size, when patch-patch interaction is not negligible. An earthquake of multiple shock type may be such a case that the size of mainshock earthquake does not necessarily scale with its nucleation zone size.

The effective λ_c can be inferred for earthquakes for which the constitutive law parameters $\Delta\tau_b$ and D_c have been estimated by assuming the magnitude for the peak shear strength τ_p appropriately (Ohnaka, 2003). Figure 13 shows how large the effective λ_c inferred for real earthquakes is. In this figure, data on small-scale frictional slip failure and fracture in the laboratory have also been plotted for comparison. From this figure, one can see that the effective λ_c for major earthquakes has a value ranging from 1 m to 100 m. One can also see from this figure that laboratory data on small-scale shear fracture and frictional slip failure, and field data on large-scale earthquakes can consistently be unified by constitutive scaling law (7), in spite of a vast scale difference between the two.

It can be seen from Fig. 13 that both D_c and λ_c are larger for larger earthquake faults. The reason for this can be summarized as follows; 1) Rupture surfaces of an inhomogeneous fault cannot be flat planes but necessarily exhibit geometric irregularity, 2) A large fault includes a geometrically large patch of high resistance to rupture growth in a statistical sense, 3) The irregular rupture surfaces of such a geometrically large patch contain a long predominant wavelength component λ_c, and 4) A large amount of D_c is required for breaking down a geometrically large patch containing large λ_c.

Figure 14 shows the physical scaling relation between the critical length L_c of the nucleation zone and the breakdown displacement D_c. In this figure, field data on large-scale earthquakes are compared with laboratory data on small-scale shear fracture and frictional slip failure. Relation (23) indicates that L_c is directly proportional to D_c, if $\Delta\tau_b$ is constant. Straight lines in Fig. 14 indicate the proportional relationships between L_c and D_c under the assumption that $\Delta\tau_b = 0.01, 0.1, 1, 10, 100$, or 1000 MPa. One can see from Fig. 14 that different sets of data on small-scale shear fracture and frictional slip failure, and large-scale earthquakes are consistently unified by theoretical scaling law (23) derived from the slip-dependent constitutive law. One can also see from Fig. 14 that the nucleation zone size scales with the breakdown displacement, though the scaling relation may severely be affected by the magnitude of the breakdown stress drop $\Delta\tau_b$.

Finally, I wish to show how long the nucleation time to the critical point is for typical, major earthquakes. Given that the effective λ_c for major earthquakes has a value ranging from 1 to 100 m (Fig. 13), the nucleation time to the critical

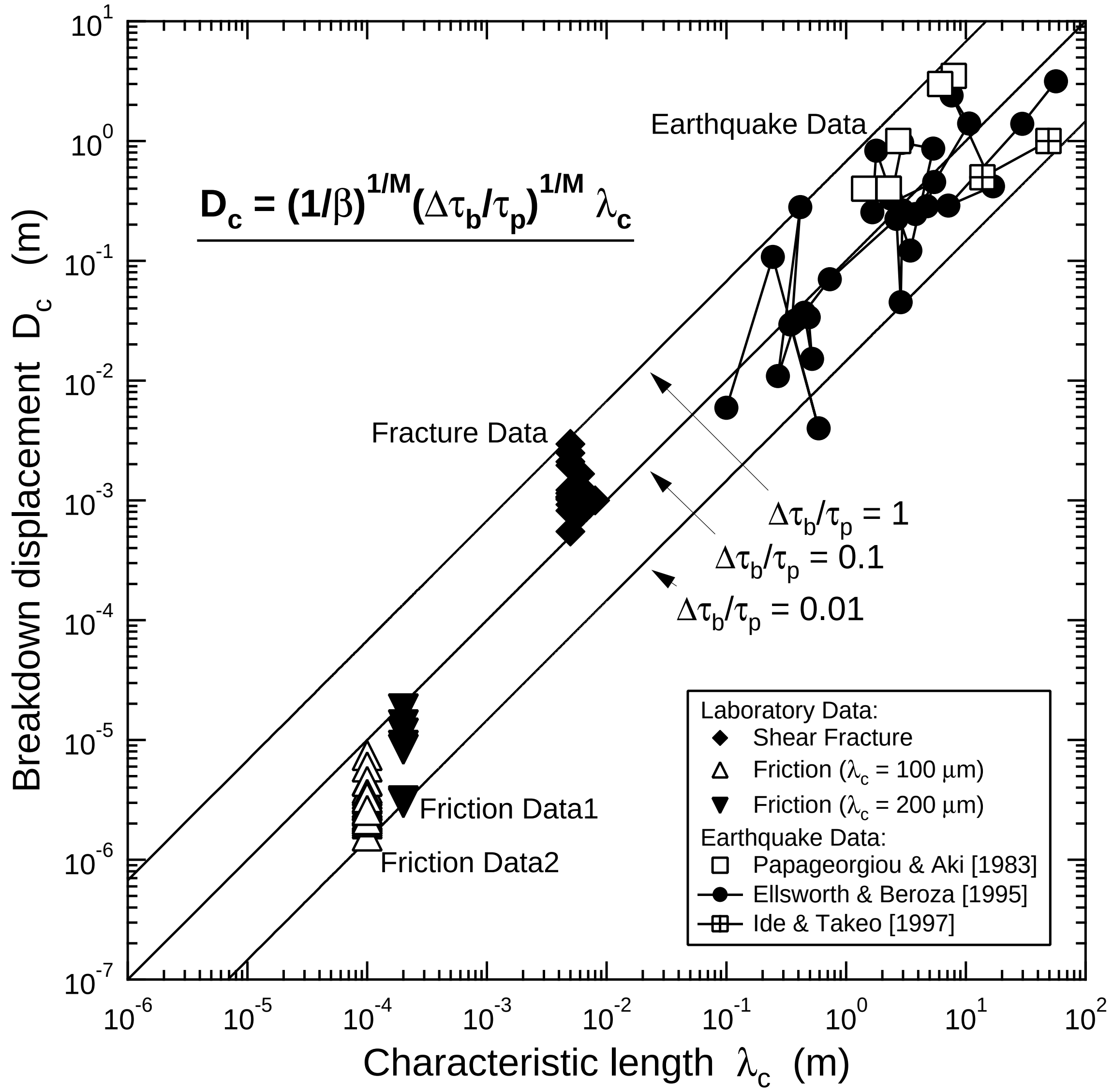

Fig. 13. Scaling relation between the breakdown displacement D_c and the characteristic length λ_c. Solid lines indicate scaling relations between D_c and λ_c when $\Delta\tau_b/\tau_p = 0.01$, 0.1 or 1 has been assumed. Field data on earthquakes and laboratory data on shear fracture and frictional slip failure are unified by scaling relation (7) in text. Reproduced from a paper by Ohnaka (2003).

point for typical earthquakes can be inferred from universal scaling relations (19) and (20), under the assumption that the exponent n is scale-invariant. Indeed, the specific time from point P in Fig. 10 to the critical point has been inferred to be of the order of several tens of minutes to a few days or much longer, depending on seismogenic environments (Ohnaka, 2004). This is contrasted with the time to the critical point for laboratory-scale rupture, which is of the order of 10 ms to 1 sec. For laboratory-scale rupture, λ_c has a value ranging from the order of 10 μm to the order of 1 mm. The nucleation time to the critical point is therefore four to five orders or much longer for typical, major earthquakes than for laboratory-scale rupture (Ohnaka, 2004).

The present results lead to the conclusion that the nucleation zone length and its duration are both longer for larger earthquakes. If the ongoing nucleation for a typical, large earthquake can be identified and monitored by any observational means, equation (17) or (19) may be useful for the short-term (immediate) forecasting (Ohnaka, 2004). Since the power law of Eq. (17) or (19) does not have any singularity over the entire time interval $t \leq t_c$, it has a great advantage, in the sense that both the occurrence time and the size of earthquake expected can be evaluated by determining t_c and L_c by curve fitting. Once L_c has been determined, the seismic moment for the earthquake expected may be inferred from scaling relation (22). This leads to the consistent conclusion that it is in principle possible to forecast a large earthquake, not only on a long-term basis (based on the seismic gap theory, see Section 3) and on an intermediate-term basis (see Section 4), but also on a short-term (or immediate) basis if the ongoing nucleation can be identified and monitored by any observational means.

7. Conclusions

A thorough discussion has been made on what the rational constitutive law for earthquake ruptures ought to be from the standpoint of the physics of rock friction and fracture on the basis of solid facts observed in the laboratory. From

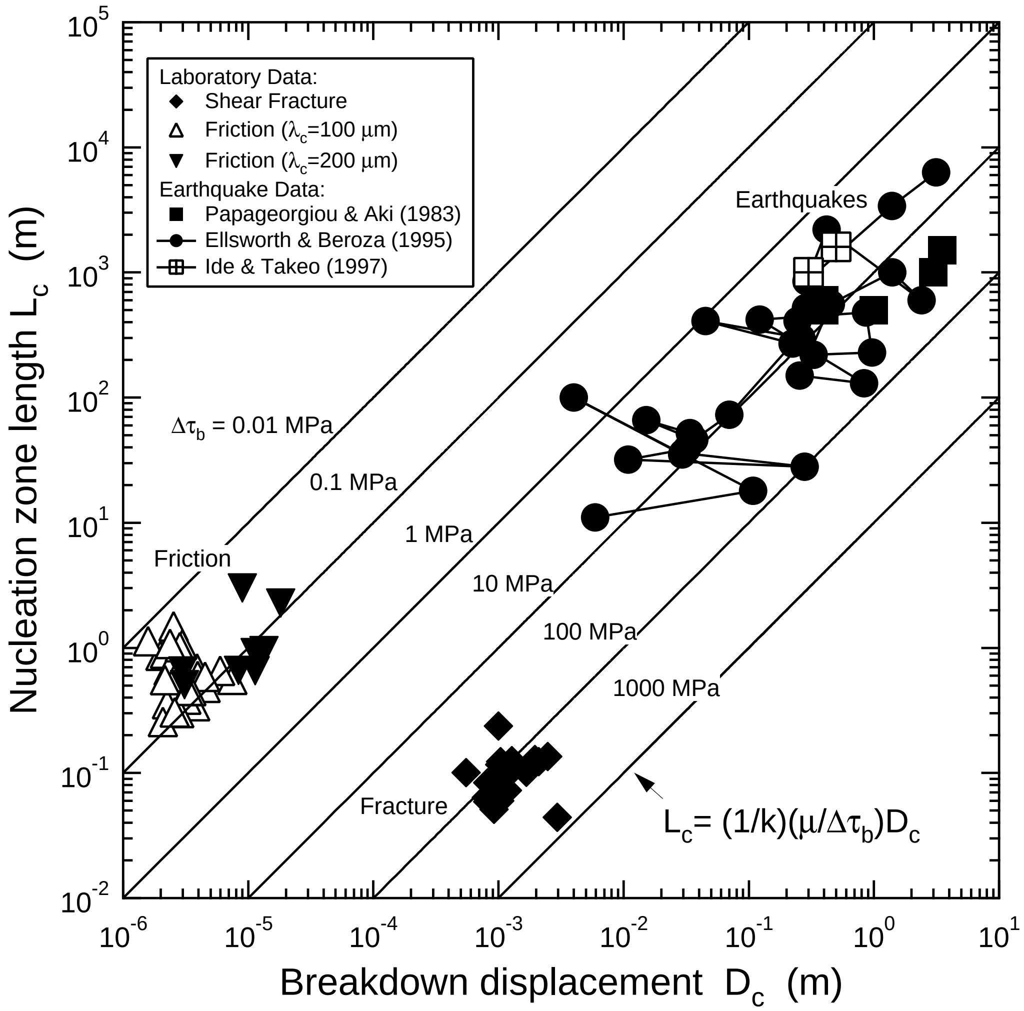

Fig. 14. Scaling relation between the nucleation zone length L_c and the breakdown displacement D_c. Contemplating that the breakdown zone length X_c is of the order of L_c, data on X_c and M_0 for earthquakes (Papageorgiou and Aki, 1983) are also plotted for comparison. Solid lines represent scaling relations between L_c and D_c when $\Delta\tau_b = 0.01, 0.1, 1, 10, 100$, or 1000 MPa has been assumed. Field data on earthquakes and laboratory data on shear fracture and frictional slip failure are unified by relation (23) in text. Reproduced from figure 16 in a paper by Ohnaka (2003).

this standpoint, it is concluded that a slip-dependent constitutive law with parameters that may be an implicit function of slip rate or time is the most rational as the governing law for earthquake ruptures. With the long-term goal of establishing a rational methodology of forecasting large earthquakes, a model of the cyclic process for a typical, large earthquake has been presented, and a comprehensive, consistent scenario for forecasting typical, large earthquakes has been discussed in terms of the underlying physics, in the context of the earthquake cycle model, by incorporating the properties of inhomogeneity and physical scaling. The entire process of one cycle for a typical, large earthquake commonly includes the following two phases: accumulation of elastic strain energy with tectonic loading (phase II), and rupture nucleation at the critical stage where an adequate amount of elastic strain energy has been accumulated (phase III). Phase II is important for physical modeling of intermediate-term forecasting, and phase III for physical modeling of short-term (immediate) forecasting. Thus, the models for intermediate-term and short-term (immediate) forecasts can be unified in the context of the earthquake cycle model. These models commonly show that there are scaling relations between the size of the ensuing event to be predicted and the critical size of the region where premonitory phenomena proceed. For typical, large earthquakes, there is also a scaling relation between the time to the critical point and the size of the event to be predicted. It follows from these scaling relations that typical, large earthquakes are in principle predictable. The scenario presented will be significant and useful as a necessary step for establishing the methodology for forecasting large earthquakes.

Acknowledgments. The original manuscript was reviewed by Raul Madariaga and an anonymous reviewer. Inspired by or in response to their comments, the paper was substantially revised to deepen the discussions of Sections 2 and 3. I am very grateful to them for their critical and thought-provoking comments. I am also very grateful to Steven McCulloch for his helpful comments, which was beneficial to improving the manuscript.

References

Aki, K., Characterization of barriers on an earthquake fault, *J. Geophys. Res.*, **84**, 6140–6148, 1979.

Aki, K., Asperities, barriers, characteristic earthquakes and strong motion prediction, *J. Geophys. Res.*, **89**, 5867–5872, 1984.

Andrews, D. J., Rupture propagation with finite stress in antiplane strain, *J.*

Geophys. Res., **81**, 3575–3582, 1976a.

Andrews, D. J., Rupture velocity of plane strain shear cracks, *J. Geophys. Res.*, **81**, 5679–5687, 1976b.

Aochi, H. and M. Matsu'ura, Slip- and time-dependent fault constitutive law and its significance in earthquake generation cycles, *Pure Appl. Geophys.*, **159**, 2029–2044, 2002.

Beroza, G. C. and T. Mikumo, Short slip duration in dynamic rupture in the presence of heterogeneous fault properties, *J. Geophys. Res.*, **101**, 22449–22460, 1996.

Bizzarri, A., M. Cocco, D. J. Andrews, and E. Boschi, Solving the dynamic rupture problem with different numerical approaches and constitutive laws, *Geophys. J. Int.*, **144**, 656–678, 2001.

Blanpied, M. L., T. E. Tullis, and J. D. Weeks, Frictional behavior of granite at low and high sliding velocities, *Geophys. Res. Lett.*, **14**, 554–557, 1987.

Bouchon, M., The state of stress on some faults of the San Andreas system as inferred from near-field strong motion data, *J. Geophys. Res.*, **102**, 11731–11744, 1997.

Bowman, D. D., G. Ouillon, C. G. Sammis, A. Sornette, and D. Sornette, An observational test of the critical earthquake concept, *J. Geophys. Res.*, **103**, 24359–24372, 1998.

Brehm, D. J. and L. W. Braile, Intermediate-term earthquake prediction using precursory events in the New Madrid seismic zone, *Bull. Seismol. Soc. Amer.*, **88**, 564–580, 1998.

Bufe, C. G. and D. J. Varnes, Predictive modeling of the seismic cycle of the greater San Francisco Bay region, *J. Geophys. Res.*, **98**, 9871–9883, 1993.

Bufe, C. G., S. P. Nishenko, and D. J. Varnes, Seismicity trends and potential for large earthquakes in the Alaska-Aleutian region, *Pure Appl. Geophys.*, **142**, 83–99, 1994.

Campillo, M. and I. R. Ionescu, Initiation of antiplane shear instability under slip dependent friction, *J. Geophys. Res.*, **102**, 20363–20371, 1997.

Campillo, M., P. Favreau, I. R. Ionescu, and C. Voisin, On the effective friction law of a heterogeneous fault, *J. Geophys. Res.*, **106**, 16307–16322, 2001.

Carlson, J. M. and J. S. Langer, Mechanical model of an earthquake fault, *Phys. Rev.*, **A 40**, 6470–6484, 1989.

Carlson, J. M., J. S. Langer, B. E. Show, and C. Tang, Intrinsic properties of a Burridge-Knopoff model of an earthquake fault, *Phys. Rev.*, **A 44**, 884–897, 1991.

Cocco, M. and A. Bizzarri, On the slip-weakening behavior of rate- and state dependent constitutive laws, *Geophys. Res. Lett.*, **29**(11), 10.1029/2001GL013999, 2002.

Day, S. M., Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress, *Bull. Seismol. Soc. Amer.*, **72**, 1881–1902, 1982.

Dieterich, J. H., Preseismic fault slip and earthquake prediction, *J. Geophys. Res.*, **83**, 3940–3948, 1978.

Dieterich, J. H., Modeling of rock friction, 1. Experimental results and constitutive equations, *J. Geophys. Res.*, **84**, 2161–2168, 1979.

Dieterich, J. H., Constitutive properties of faults with simulated gouge, in *Mechanical Behavior of Crystal Rocks,* edited by N. L. Cater, M. Friedman, J. M. Logan, and D. W. Stearns, Geophysical Monograph, 24, pp. 103–120, American Geophysical Union, Washington, D.C., 1981.

Dieterich, J. H., A model for the nucleation of earthquake slip, in *Earthquake Source Mechanics,* edited by S. Das, J. Boatwright, and C. H. Scholz, pp. 37–47, Geophysical Monograph, 37, American Geophysical Union, Washington, D.C., 1986.

Dieterich, J. H. and B. Kilgore, Implications of fault constitutive properties for earthquake prediction, *Proc. Natl. Acad. Sci. USA*, **93**, 3787–3794, 1996.

Dodge, D. A., G. C. Beroza, and W. L. Ellsworth, Evolution of the 1992 Landers, California, foreshock sequence and its implications for earthquake nucleation, *J. Geophys. Res.*, **100**, 9865–9880, 1995.

Dodge, D. A., G. C. Beroza, and W. L. Ellsworth, Detailed observations of California foreshock sequence: Implications for the earthquake initiation process, *J. Geophys. Res.*, **101**, 22371–22392, 1996.

Ellsworth, W. L. and G. C. Beroza, Seismic evidence for an earthquake nucleation phase, *Science,* **268**, 851–855, 1995.

Engdahl, E. R. and A. Villasenor, Global seismicity: 1900–1999, in *International Handbook of Earthquake and Engineering Seismology, Part A,* edited by W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, pp. 665–690, Academic Press, New York, 2002.

Fedotov, S. A., Regularities of the distribution of strong earthquakes in Kamchatka, the Kurile Islands and northeastern Japan, *Acad. Sci. USSR Trudy Inst. Phys. Earth,* **36**, 66–93, 1965.

Fukuyama, E. and R. Madariaga, Dynamic propagation and interaction of a rupture front on a planer fault, *Pure Appl. Geophys.*, **157**, 1959–1979, 2000.

Fukuyama, E. and K. B. Olsen, A condition for super-shear rupture propagation in a heterogeneous stress field, *Pure Appl. Geophys.*, **159**, 2047–2056, 2002.

Gu, J.-C., J. R. Rice, A. L. Ruina, and S. T. Tse, Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction, *J. Mech. Phys. Solids,* **32**, 167–196, 1984.

Ida, Y., Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy, *J. Geophys. Res.*, **77**, 3796–3805, 1972.

Ida, Y., The maximum acceleration of seismic ground motion, *Bull. Seismol. Soc. Amer.*, **63**, 959–968, 1973.

Ide, S. and M. Takeo, Determination of constitutive relations of fault slip based on seismic wave analysis, *J. Geophys. Res.*, **102**, 27379–27391, 1997.

Imamura, A., On the seismic activity of central Japan, *Japanese Journal of Astronomy and Geophysics,* Transactions, National Research Council of Japan, **6**, 119–137, 1928/29.

Ionescu, I. R. and M. Campillo, Influence of the shape of the friction law and fault finiteness on the duration of initiation, *J. Geophys. Res.*, **104**, 3013–3024, 1999.

Ishibashi, K. and K. Satake, Problems on forecasting great earthquakes in the subduction zones around Japan by means of paleoseismology, *J. Seismol. Soc. Japan,* Second Series, **50** (Supplement), 1–21, 1998.

Jaume, S. C. and L. R. Sykes, Evolving towards a critical point: A review of accelerating seismic moment/energy release prior to large and great earthquakes, *Pure Appl. Geophys.*, **155**, 279–306, 1999.

Kanamori, H., The nature of seismicity patterns before large earthquakes, in *Earthquake Prediction—An International Review,* edited by D. W. Simpson and P. G. Richards, Maurice Ewing Series 4, pp. 1–19, American Geophysical Union, Washington D.C., 1981.

Kanamori, H., Earthquake prediction: An overview, in *International Handbook of Earthquake and Engineering Seismology, Part B,* edited by W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, pp. 1205–1216, Academic Press, New York, 2002.

Kanamori, H. and G. S. Stewart, Seismological aspects of the Guatemala earthquake of February 4, 1976, *J. Geophys. Res.*, **83**, 3427–3434, 1978.

Kato, A., M. Ohnaka, and H. Mochizuki, Constitutive properties for the shear failure of intact granite in seismogenic environments, *J. Geophys. Res.*, **108**(B1), 2060, doi: 10.1029/2001JB000791, 2003a.

Kato, A., M. Ohnaka, S. Yoshida, and H. Mochizuki, Effects of strain rate on constitutive properties for the shear failure of intact granite in seismogenic environments, *Geophys. Res. Lett.*, **30**(21), 2108, doi:10.1029/2003GL018372, 2003b.

Kumamoto, T., Long-term conditional seismic hazard of Quaternary active faults in Japan, *J. Seismol. Soc. Japan,* Second Series, **50** (Supplement), 53–71, 1998.

Linker, M. F. and J. H. Dieterich, Effects of variable normal stress on rock friction: Observations and constitutive equations, *J. Geophys. Res.*, **97**, 4923–4940, 1992.

Madariaga, R. and K. B. Olsen, Criticality of rupture dynamics in 3-D, *Pure Appl. Geophys.*, **157**, 1981–2001, 2000.

Madariaga, R. and K. B. Olsen, Earthquake dynamics, in International Handbook of Earthquake and Engineering Seismology, Part A, edited by W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, pp. 175–194, Academic Press, New York, 2002.

Madariaga, R., K. B. Olsen, and R. J. Archuleta, Modeling dynamic rupture in a 3-D earthquake fault model, *Bull. Seismol. Soc. Amer.*, **88**, 1182–1197, 1998.

Maeda, K., Time distribution of immediate foreshocks obtained by a stacking method, *Pure Appl. Geophys.*, **155**, 381–394, 1999.

Masuda, K., H. Mizutani, and I. Yamada, Experimental study of strain-rate dependence and pressure dependence of failure properties of granite, *J. Phys. Earth,* **35**, 37–66, 1987.

Matsuda, T., Present state of long-term prediction of earthquakes based on active fault data in Japan—An example for the Itoigawa-Shizuoka tectonic line active fault system—, *J. Seismol. Soc. Japan,* Second Series, **50** (Supplement), 23–33, 1998.

Matsu'ura, M., H. Kataoka, and B. Shibazaki, Slip-dependent friction law and nucleation processes in earthquake rupture, *Tectonophysics,* **211**, 135–148, 1992.

McCann, W. R., S. P. Nishenko, L. R. Sykes, and J. Krause, Seismic gaps and plate tectonics: Seismic potential for major boundaries, *Pure Appl. Geophys.*, **117**, 1082–1147, 1979.

Mogi, K., Sequential occurrence of recent great earthquakes, *J. Phys. Earth,* **16**, 30–36, 1968.

Nakanishi, H., Earthquake dynamics driven by a viscous fluid, *Phys. Rev.*, **A 46**, 4689–4692, 1992.

Nishenko, S. P., Circum-Pacific seismic potential: 1989–1999, *Pure Appl. Geophys.*, **135**, 169–259, 1991.

Ohnaka, M., Earthquake source nucleation: a physical model for short-term precursors, *Tectonophysics,* **211**, 149–178, 1992.

Ohnaka, M., Critical size of the nucleation zone of earthquake rupture inferred from immediate foreshock activity, *J. Phys. Earth,* **41**, 45–56, 1993.

Ohnaka, M., A shear failure strength law of rock in the brittle-plastic transition regime, *Geophys. Res. Lett.*, **22**, 25–28, 1995.

Ohnaka, M., Nonuniformity of the constitutive law parameters for shear rupture and quasistatic nucleation to dynamic rupture: A physical model of earthquake generation processes, *Proc. Natl. Acad. Sci. USA*, **93**, 3795–3802, 1996.

Ohnaka, M., Earthquake generation processes and earthquake prediction: Implications of the underlying physical law and seismogenic environments, *J. Seismol. Soc. Japan,* Second Series, **50** (Supplement), 129–155, 1998.

Ohnaka, M., A physical scaling relation between the size of an earthquake and its nucleation zone size, *Pure Appl. Geophys.*, **157**, 2259–2282, 2000.

Ohnaka, M., A constitutive scaling law and a unified comprehension for frictional slip failure, shear fracture of intact rock, and earthquake rupture, *J. Geophys. Res.*, **108**(B2), 2080, doi: 10.1029/2000JB000123, 2003.

Ohnaka, M., A constitutive scaling law for shear rupture that is inherently scale-dependent, and physical scaling of nucleation time to critical point, *Pure Appl. Geophys.*, 2004 (in press).

Ohnaka, M. and Y. Kuwahara, Characteristic features of local breakdown near a crack-tip in the transition zone from nucleation to unstable rupture during stick-slip shear failure, *Tectonophysics,* **175**, 197–220, 1990.

Ohnaka, M. and L.-f. Shen, Scaling of the shear rupture process from nucleation to dynamic propagation: Implications of geometric irregularity of the rupturing surfaces, *J. Geophys. Res.*, **104**, 817–844, 1999.

Ohnaka, M. and T. Yamashita, A cohesive zone model for dynamic shear faulting based on experimentally inferred constitutive relation and strong motion source parameters, *J. Geophys. Res.*, **94**, 4089–4104, 1989.

Ohnaka, M., Y. Kuwahara, and K. Yamamoto, Constitutive relations between dynamic physical parameters near a tip of the propagating slip zone during stick-slip shear failure, *Tectonophysics,* **144**, 109–125, 1987.

Ohnaka, M., Y. Kuwahara, K. Yamamoto, and T. Hirasawa, Dynamic breakdown processes and the generating mechanism for high-frequency elastic radiation during stick-slip instabilities, in *Earthquake Source Mechanics,* edited by S. Das, J. Boatwright, and C. H. Scholz, Geophysical Monograph, 37, pp. 13–24, American Geophysical Union, Washington, D.C., 1986.

Ohnaka, M., S. Yoshida, L.-f. Shen, and H. Mochizuki, Slip-failure nucleation processes and microseismicity, Seismol. Soc. Japan, Programme and Abstracts, Fall Meeting, No. 2, p. 298, 1993.

Ohnaka, M., M. Akatsu, H. Mochizuki, A. Odedra, F. Tagashira, and Y. Yamamoto, A constitutive law for the shear failure of rock under lithospheric conditions, *Tectonophysics*, **277**, 1–27, 1997.

Okubo, P. G., Dynamic rupture modeling with laboratory-derived constitutive relations, *J. Geophys. Res.*, **94**, 12321–12335, 1989.

Okubo, P. G. and J. H. Dieterich, State variable fault constitutive relations for dynamic slip, in *Earthquake Source Dynamics,* edited by S. Das, J. Boatwright, and C. H. Scholz, Geophysical Monograph, 37, pp. 25–35, American Geophysical Union, Washington, D.C., 1986.

Palmer, A. C. and J. R. Rice, The growth of slip surfaces in the progressive failure of over-consolidated clay, *Proc. Roy. Soc. London*, **A 332**, 527–548, 1973.

Papageorgiou, A. S. and K. Aki, A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion. Part II. Applications of the model, *Bull. Seismol. Soc. Amer.*, **73**, 953–978, 1983.

Rice, J. R., The mechanics of earthquake rupture, in *Physics of the Earth's Interior,* edited by A. M. Dziewonski and E. Boschi, pp. 555–649, North-Holland, Amsterdam, 1980.

Rice, J. R., Constitutive relations for fault slip and earthquake instabilities, *Pure Appl. Geophys.*, **121**, 443–475, 1983.

Rice, J. R., Shear instability in relation to the constitutive description of fault slip, *Proceedings of the 1st International Congress on Rockbursts and Seismicity in Mines,* pp. 57–62, Johannesburg, 1984.

Rice, J. R. and A. L. Ruina, Stability of steady frictional slipping, *J. Appl. Mech.*, Trans. ASME, **50**, 343–349, 1983.

Rudnicki, J. W., Fracture mechanics applied to the Earth's crust, *Ann. Rev. Earth Planet. Sci.*, **8**, 489–525, 1980.

Rudnicki, J. W., Physical models of earthquake instability and precursory processes, *Pure Appl. Geophys.*, **126**, 531–554, 1988.

Ruina, A., Slip instability and state variable friction laws, *J. Geophys. Res.*, **88**, 10359–10370, 1983.

Ruina, A. L., Constitutive relations for frictional slip, in *Mechanics of Geomaterials,* edited by Z. Bazant, pp. 169–188, John Wiley, New York, 1985.

Rundle, J. B., W. Klein, D. L. Turcotte, and B. D. Malamud, Precursory seismic activation and critical-point phenomena, *Pure Appl. Geophys.*, **157**, 2165–2182, 2000.

Saito, M., Forecasting time of slope failure by tertiary creep, *Proceedings of 7th International Conference on Soil Mechanics and Foundation Engineering,* **2**, 677–683, 1969.

Schwartz, D. P. and K. J. Coppersmith, Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res.*, **89**, 5681–5698, 1984.

Shibazaki, B. and M. Matsu'ura, Spontaneous processes for nucleation, dynamic propagation, and stop of earthquake rupture, *Geophys. Res. Lett.*, **19**, 1189–1192, 1992.

Shibazaki, B. and M. Matsu'ura, Foreshocks and pre-events associated with the nucleation of large earthquakes, *Geophys. Res. Lett.*, **22**, 1305–1308, 1995.

Shibazaki, B. and M. Matsu'ura, Transition process from nucleation to high-speed rupture propagation: Scaling from stick-slip experiments to natural earthquakes, *Geophys. J. Int.*, **132**, 14–30, 1998.

Sieh, K., The repetition of large-earthquake ruptures, *Proc. Natl. Acad. Sci. USA*, **93**, 3764–3771, 1996.

Sykes, L. R. and S. P. Nishenko, Probabilities of occurrence of large plate rupturing earthquakes for the San Andreas, San Jacinto, and Imperial faults, California, 1983–2003, *J. Geophys. Res.*, **89**, 5905–5927, 1984.

Tanaka, S., M. Ohtake, and H. Sato, Evidence for tidal triggering of earthquakes as revealed from statistical analysis of global data, *J. Geophys. Res.*, **107**(B10), 2211, doi:10.1029/ 2001JB001577, 2002.

Tullis, T. E. and J. D. Week, Constitutive behavior and stability of frictional sliding of granite, *Pure Appl. Geophys.*, **124**, 383–414, 1986.

Uenishi, K. and J. R. Rice, Universal nucleation length for slip-weakening rupture instability under nonuniform fault loading, *J. Geophys. Res.*, **108**(B01), 2042, doi: 10.1029/2001JB001681, 2003.

Utsu, T., Estimation of parameters for recurrence models of earthquakes, *Bull. Earthquake Res. Inst., University of Tokyo,* **59**, 53–66, 1984.

Utsu, T., Seismicity patterns and long-term prediction of large earthquakes—Seismic cycles, gaps, quiescence, precursory activities, migration, correlation, etc.—, *J. Seismol. Soc. Japan,* Second Series, **50** (Supplement), 73–82, 1998.

Utsu, T., *Seismicity Studies: A Comprehensive Review,* 876 pp., University of Tokyo Press, Tokyo, 1999.

Utsu, T., Statistical features of seismicity, in *International Handbook of Earthquake and Engineering Seismology, Part A,* edited by W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, pp. 719–732, Academic Press, New York, 2002.

Varnes, D. J., Predicting earthquakes by analyzing accelerating precursory seismic activity, *Pure Appl. Geophys.*, **130**, 661–686, 1989.

Yamashita, T. and M. Ohnaka, Nucleation process of unstable rupture in the brittle regime: A theoretical approach based on experimentally inferred relations, *J. Geophys. Res.*, **96**, 8351–8367, 1991.

Yin, X.-C., X.-Z. Chen, Z.-P. Song, and C. Yin, A new approach to earthquake prediction: The load/unload response ratio (LURR) theory, *Pure Appl. Geophys.*, **145**, 701–715, 1995.

Yin, X.-C., Y.-C. Wang, K.-Y. Peng, Y.-L. Bai, H.-T. Wang, and X.-F. Yin, Development of a new approach to earthquake prediction: Load/unload response ratio (LURR) theory, *Pure Appl. Geophys.*, **157**, 2365–2383, 2000.

Yin, X.-C., P. Mora, K. Peng, Y. C. Wang, and D. Weatherley, Load-unload response ratio and accelerating moment/energy release critical region scaling and earthquake prediction, *Pure Appl. Geophys.*, **159**, 2511–2523, 2002.

M. Ohnaka (e-mail: ohnaka@g05.itscom.net)