Jour. Fac. Sci., Hokkaido Univ., Ser. VII (Geophysics), Vol. 11, No. 3, 665-690, 1999.

Stochastic Scaling and Non-linear Scaling Law of Complex Earthquake Activity

Junji Koyama

Division of Earth and Planetary Sciences, Graduate School of Science Hokkaido University, Sapporo, 060-0810 Japan

Yoshinobu Motoya

Research Center for Earthquake and Volcano, Graduate School of Science Hokkaido University, Sapporo, 060-0810 Japan

and

Hiroaki Hara

Graduate School of Engineering, Tohoku University, Sendai, 980-77 Japan

(Received November 30, 1998)

Abstract

The randomness of earthquake occurrences is investigated in order to elucidate their universal nature as a complex system. The essential difference of this complex earthquake activity from the physical complex system is its size-effect in wide range. One series of an earthquake activity can be selected by taking a particular value of earthquake magnitude. This series of earthquakes is described by the same equation of time evolution characterizing a particular stochastic process. Taking another magnitude value, another series of the earthquake activity can also be selected, which is described by a different equation representing another stochastic process. Many of these series of the earthquake activities form a cluster of earthquake occurrences with a wide magnitude range. The stochastic property among the different series of earthquakes is characterized by a stochastic scaling specifying a scaling relation among different random processes (series of earthquakes). The scaling parameters can be determined by the maximum entropy condition. In order to unify statistical properties of local, regional and global earthquake activities, a non-linear scaling law is derived. The non-linear scaling law characterizes the hierarchy of the complexity for the composite system consisting of component complex systems in a general manner.

1. Introduction

The earthquake source is represented by a complex faulting process composed of random ruptures of small-scale fault heterogeneities. The fluctuation of stresses on a fault plane is very important for the generation of an *isolated* earthquake, mainshock-aftershocks and an earthquake swarm (Cheng and Knopoff, 1987). The spatial distribution of these fault heterogeneities relates to locked and unlocked fault patches (fault asperities and barriers) on the heterogeneous fault plane. These fault patches appear to exist on all scales producing stress concentration locally on the fault (e.g., Koyama, 1994; Ruff, 1992; Yamashita and Knopoff, 1992).

The stress concentration around fault patches becomes the source area of aftershocks. Aftershocks following large shallow earthquakes are considered to be the relaxation of such stress concentrations induced by a particular mainshock. The time evolution of this relaxation process is described empirically by Omori-Utsu's formula as a power law of time;

$$n(t) \propto t^{-p} \tag{1}$$

where n is the number of aftershocks at a time t after the origin of the mainshock. Power constant p has been determined by a more or less subjective manner to be about 0.7 to 1.4 for large earthquakes in the world (Utsu, 1969; Wang, 1994).

Many models have been proposed to explain Omori-Utsu's empirical formula. Making simple assumptions on the earthquake sequence, a logistic growth of aftershocks has been investigated by Ouchi (1993). Some introduced a non-elastic process on the fault to explain the time delays associated with aftershocks: visco-elastic friction by Burridge and Knopoff (1967), and stress corrosion cracking by Das and Scholz (1981) and Yamashita and Knopoff (1987, 1992). Creep and stress relaxations on the fault are another consideration on the time delay of aftershocks (Gu et al., 1979). These models predict the value of p of about 1.0.

The power law of time is also applied for earthquake swarms and volcanic earthquakes. The empirical constant of p is somewhat larger in these cases, about 1.6 to 2, than those for aftershocks. An example is shown in Fig. 1 for the earthquake swarm in Izu peninsula, Japan, for which the value of p is about 1.6. The value is too large to be reasonably explained by the previous models listed above. Omori-Utsu's empirical formula is different from the simple exponential decay by Poisson process, i.e., the formula is representing one aspect of the complexity of earthquake phenomena and the long-tail behavior of earthquake activity.

In case of earthquake swarms the number of earthquakes increases initially

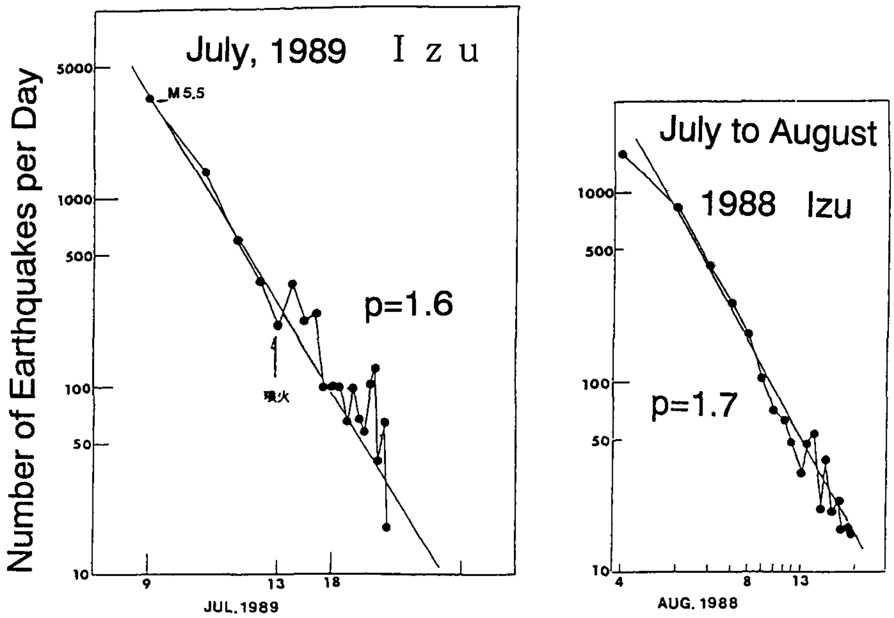


Fig. 1. Time-decay of Izu earthquake swarm, Japan in July of 1989 and in July to August of 1988 (Tokyo University, 1989). Hyperbolic decay rate p of earthquake number is about 1.6 on the left and 1.7 on the right, respectively.

and then decreases gradually with time. An exponential time-decay was seen for the Matsushiro earthquake swarm activity in Japan (Mogi, 1991). Exponential time-decay has also been found for earthquake swarms in East-off Izu peninsula, Japan (Tsukuda, 1993). An exponential time-decay indicates that such activity is a Poisson process. We do not know any physical relationship between the exponential time-decay of earthquake swarm events and the power-law decay of aftershocks, although some studies derive such a relation based on physically-unproved a priori assumption.

The Poisson process indicates the probability density of time intervals of events $P(\tau)=\lambda\ \exp(-\lambda|\tau|)$, where τ is a time lag, and λ is an average number of events in unit of time. This exponential behavior of earthquake occurrences has been found for the largest shallow earthquakes in and near Japan and for the large deep earthquakes in the world (Figs. 2 and 3). Therefore, the occurrence of the largest earthquakes in the world is believed to be a purely random phenomenon. Meanwhile, the autocorrelation function of the global earthquake activity is not like an exponential one but like a power law of $|\tau|^{2H-2}$, where H is an empirical constant of $0<H\leq 1$ (Ogata and Abe, 1991).

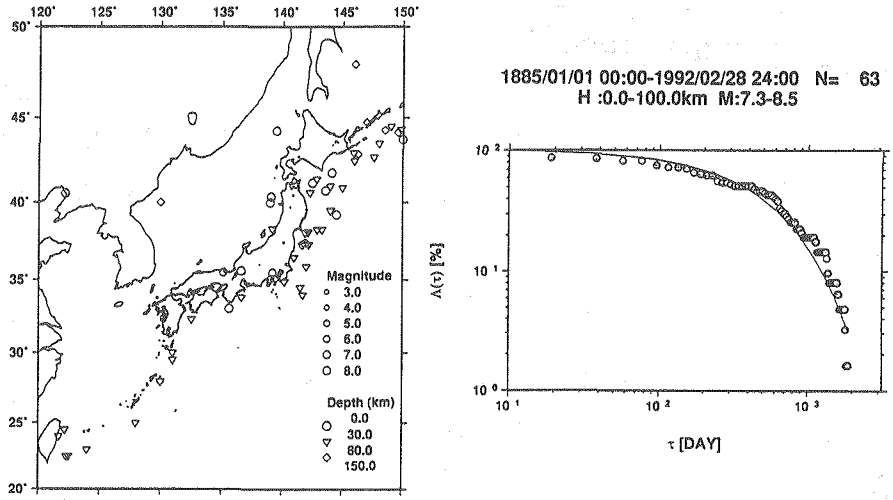


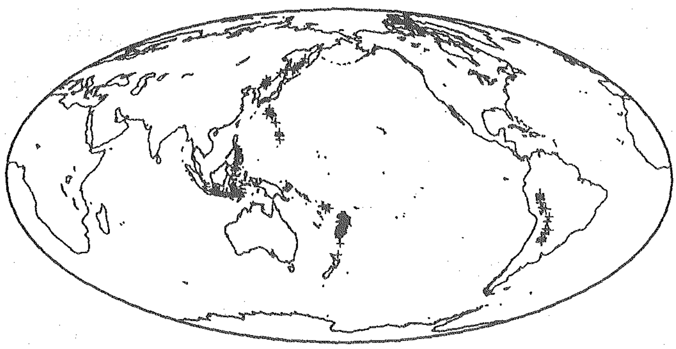
Fig. 2. Cumulative number of time-intervals of successive earthquakes in and near Japan. Large and great shallow earthquakes from 1885 to 1992 (earthquake magnitudes larger than or equal to 7.3) are analyzed. Epicenters are plotted by different symbols representing focal depths.

Fault systems and joints in rock specimens are understood as a fractal (Brown and Scholz, 1985 ; Turcotte, 1997) represented by a power-law function. Therefore, the size-number distribution of random fault patches is also considered to be of a power law. The fractal nature of earthquake occurrences is also known in seismology as the Gutenberg and Richter's (hereafter abbreviated as G-R) relation (1954) :

$$\log N(M_s) = a - bM_s, \tag{2}$$

where N is the number of earthquakes larger than M_s in a certain area, or the number of earthquakes in a surface-wave magnitude range from M_s to $M_s + \Delta M_s$. Parameters of a and b are positive constants. It indicates the power-law distribution of number and strength of earthquake sources, since the magnitude relates to the logarithm of earthquake energy (Gutenberg and Richter, 1954 ; Koyama, 1994). The G-R relation is valid not only for natural earthquakes in the world but also for aftershocks of large earthquakes in a particular region. The b-value is almost constant of about 1 irrespectively to the size and to seismic region of earthquakes.

The Goishi model of Ohtsuka (1972) and the branching model of Vera-Jones (1976) are purely stochastic, and provide a similar magnitude-frequency relation



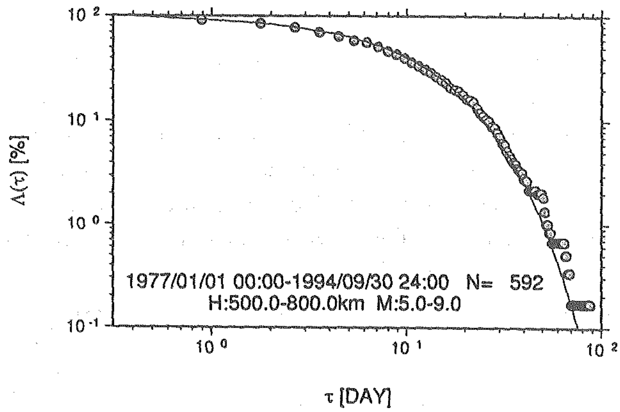


Fig. 3. Cumulative number of time-intervals of successive deep-focus earthquakes worldwide. Earthquakes from 1977 to 1994 are tabulated by the U.S. Geological Survey. Magnitudes are larger than 5.5 and focal depths more than 500 km.

to the G-R relation. The Goishi model is not based on the dynamics of rupture propagation on a fault plane but on a probabilistic growth like a tree-like shape or diffusion-limited-aggregation. Self-organized criticality is a recent idea for understanding complicated earthquake activity (Bak and Tang, 1989; Carlson, 1991). These numerical simulations indicate that the earthquake is a random phenomena with a small number of freedoms under a specific simple rule and that the earthquake as a complex system is in the critical state of phase transitions. Therefore, it is tempting to search for a universality in the fundamental physical origin of complex earthquake activity in relation to the complex system in physics.

2. Stochastic scaling of earthquake activity

There occur a number of aftershocks following each mainshock of large earthquake. Aftershocks distribute on a fault plane of the mainshock more or less in a random manner. Studies on the statistical nature of earthquakes have tested for a Poisson and for a Markov process (Aki, 1956 ; Wu, 1990). These studies view the earthquake source as a point without any characteristic energy nor finite source size. This viewpoint of an earthquake statistics has been applied to the recent studies (e.g., Ogata and Abe, 1991).

The number density of aftershocks decreases systematically as a power law described by (1) and does not decrease in a statistically homogeneous manner. Furthermore, the size of earthquakes varies by many orders of magnitudes. Some are large enough to generate secondary disastrous earthquakes and some are so small that can be detected only by sensitive instruments. The randomness of occurrence times and the size distribution are inherent to the earthquake properties, and must be considered in order to obtain a complete understanding of the earthquake statistics. This attitude is quite different from previous studies.

The number $n(t)$ of aftershocks is the sum of aftershocks with different magnitude ranges as

$$n(t)=\sum_{i=0}^{M} n_i(t), \tag{3}$$

where $n_i(t)$ represents one series of aftershocks classified by an i-th magnitude range in a time interval $t \sim t+dt$. The magnitude range specified by $i=0$ is the minimum magnitude observed in a particular earthquake activity and that by $i=M$ indicates that with the maximum magnitude. Because of the randomness, $n_i(t)$ and $n_j(t)$ are independent, provided that $i \neq j$. This type of size-dependent randomness has been neglected, although it is very important to account for the scaling and the energy of the earthquake activity. Suppose that there are left N_i^0 nuclei produced by the mainshock for succeeding aftershocks of the i-th magnitude range. The number of aftershocks of the i-th magnitude range within the time interval, $n_i(t)dt$, can be written as

$$n_i(t)dt=-dN_i(t)=\mu_i(t)N_i(t)dt, \tag{4}$$

where $N_i(t)$ is the number of nuclei classified by the i-th magnitude range to be ruptured after the time t and $\mu_i(t)dt$ is the probability of rupturing one nucleus at the time interval. The initial condition for (4) has been given by N_i^0 at $t=0$. In general, there may be a weak time-dependency and also magnitude dependence of $\mu_i(t)$. Some studies a priori assume such dependence without any physical justification. However, we assume $\mu_i(t)$ to be a constant of μ_0 throughout this study for the sake of simplicity.

We may not be able to characterize the stochastic size-dependence of earthquake activity by such a small number of macroscopic parameters. We consider that each series of activity is governed by the same equation of temporal evolution in (4) with respective stochastic behavior. A mathematically simple assumption is to derive a respective scaling behavior for N_i's, though they are random functions and statistically independent. A stochastic scaling has been introduced to specify the scaling relation for such random functions by Koyama and Hara (1993):

$$f_d : \frac{\alpha_d}{\beta_d} N_i(\beta_d t) \rightarrow N_{i+1}(t), \tag{5}$$

where α_d and β_d are scaling parameters, positive constants, and both are smaller than unity. When $\alpha_d = \beta_d$, the stochastic scaling is self-similar. When $\alpha_d \neq \beta_d$, it is self-affine. However, this scaling of f_d does not specify the self-similarity and/or the self-affinity of functional forms, but represents a mapping which describes the similarity nature underlying statistical properties of N_i's. Scaling in a fractal geometry and in a scale invariant nature introduces a relation of some function $f(x)$ as $f(ax) = a^H f(x)$ (Feder, 1989; Koyama and Feng, 1995), where H and a are positive constants. The stochastic scaling in this study specifies a scaling relation of statistical moments and/or autocorrelation of random functions, such as $N_i(t)$'s. Therefore, the present scaling is designated as *stochastic*.

The solution of (3) and (4) is now straightforward under the stochastic scaling of (5) and it describes the number of aftershocks at a time of t since the mainshock as

$$n(t) = -\frac{d}{dt}\left[\sum_{i=0}^{M} N_0^0 \left(\frac{\alpha_d}{\beta_d}\right)^i \exp(-\mu_0 \beta_d^i t)\right]. \tag{6}$$

When $0 \leq \mu_0 t \ll 1$, (6) is related to the assumed initial condition. When $M \gg 1$ and $t \gg 1/\mu_0$, an asymptotic form of the solution can be obtained by the aid of the steepest descent method (Koyama and Hara, 1992). Then the solution of (6) is rewritten

$$n(t) \simeq A_d (\mu_0 t)^{-\xi_d}, \tag{7}$$

$$A_d = \mu_0 N_0^0 \sqrt{\frac{\pi}{2 \ln(\alpha_d/\beta_d) \ln \beta_d}} (\xi_d - 1) \exp([\xi_d - 1]\,[\ln(\xi_d - 1) - 1]), \tag{8}$$

where ξ_d is a fractal dimension defined by the scaling parameters as

$$\xi_d = \frac{\ln \alpha_d}{\ln \beta_d}. \tag{9}$$

The asymptotic form (7) is also valid for both α_d and β_d larger than unity. The solution (6) indicates the exponential decay of an earthquake activity, whereas its asymptotic form (7) manifests the power-law decay. Therefore, these solutions surely give an insight into the Omori-Utsu's empirical decay of aftershocks and the exponential decay of earthquake swarms. A relationship between the Omori-Utsu's empirical constant p and the fractal dimension of the stochastic scaling is derived as

$$p = \xi_d. \tag{10}$$

This relationship points out the physical meaning of the empirical power coefficient p in terms of the scaling parameters of complex earthquake activity.

3. Long-tail behavior of earthquake activity

3.1 Observations

The statistical behavior of earthquakes has been investigated for many years. For many cases of great earthquakes a Poisson distribution fits the observed frequency. In this case, the earthquakes are mutually independent. Since such behavior is found for the data of about 100 years of modern seismometry, the probability of earthquake occurrence is considered to be independent of time and of total number of events. Fig. 2 reconfirms this Poisson distribution of time-intervals of successive great shallow-earthquakes in and near Japan. Fig. 3 shows the same plot for large deep-focus earthquakes worldwide. Poisson distributions, drawn solid curves in the figures, explain the observed data, demonstrating the randomness of large earthquakes.

Figure 4 shows another example of the probability estimated from the cumulative number of time-intervals of regional earthquakes in and near Japan. Because of the historical changes in earthquake detectability, the minimum magnitude of 5.0 (Japan Meteorological Agency scale) and focal depths shallower than or equal to 60 km are used to prepare the database for analysis. The result is that the distribution for a particular earthquake activity classified by a narrow magnitude range can be described by a Poisson distribution. A distribution similar to the Poisson can be found not only for the small magnitude

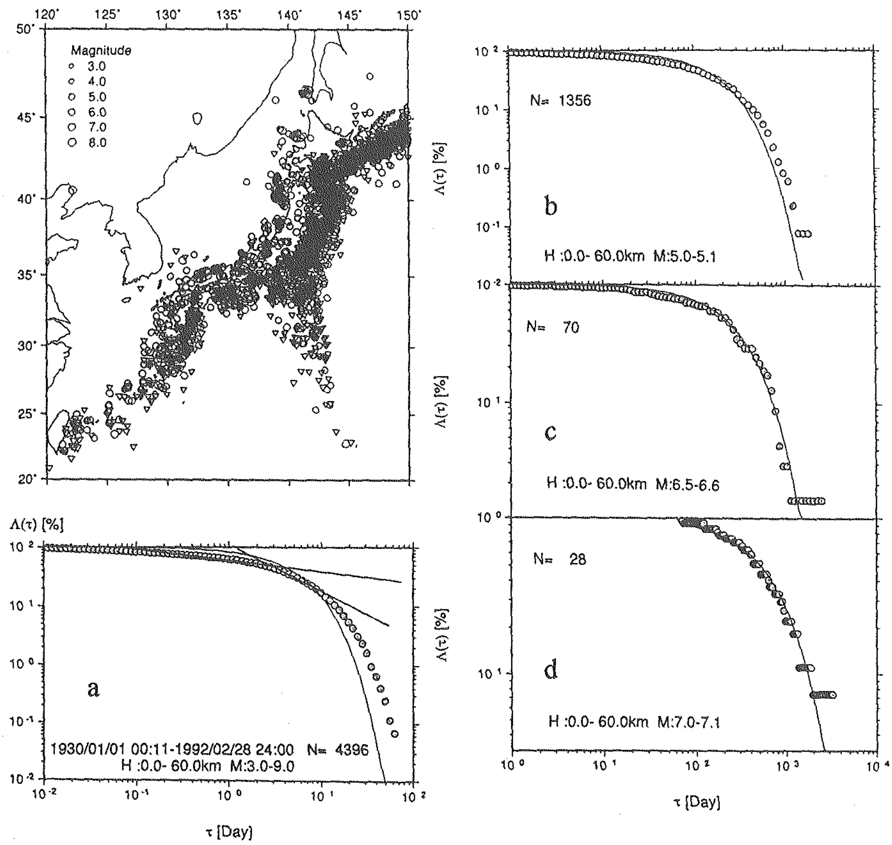


Fig. 4. Cumulative number of time intervals of successive shallow earthquakes in Japan and vicinity. Number of earthquakes N is normalized to represent the probability distribution in %. The earthquakes are from the Japan Meteorological Agency catalog for the period from January 1st, 1930 to February 28th, 1992 and have focal depths less than or equal to 60 km. (a) All earthquakes in the analysis. The curve in the figure is the Poisson probability distribution calculated from the mean number of earthquakes. It does not fit the observed time-interval distribution. Two straight lines are for a power-law dependence between time interval and cumulative number in days. (b) Similar plot for earthquakes in the magnitude range from 5.0 to 5.1. (c) Earthquakes in the magnitude range from 6.5 to 6.6. (d) Earthquakes in the magnitude range from 7.0 to 7.1.

range of about 5 but also for the large magnitude range of about 8. This agrees with the result in Figs. 2 and 3 for the great earthquakes. However, the time-interval distribution of all the magnitude ranges of earthquakes is no longer consistent with a Poisson distribution. The decay rates of small and medium as well as very large time-intervals are very small compared with the Poisson. We find a Poisson-like distribution for each earthquake activity classified by each narrow magnitude range, in which only the mean value of earthquake occurrences characterizes the probability distribution. However, a convex distribution is found for the cluster of all earthquakes.

Evidence for a long-range statistical dependence of earthquake activities has been presented by Ogata and Abe (1991) by investigating the autocorrelation of the global earthquake activity. They showed this long-tail behavior represented by a power law of $|\tau|^{2H-2}$ where τ is the time lag and H is an empirical constant of about 0.6 (statistically $0<H\leq 1$). They also found the same power law for regional large and great earthquakes in and near Japan, where again H is also about 0.6, although the two data sets differ in magnitude and number of

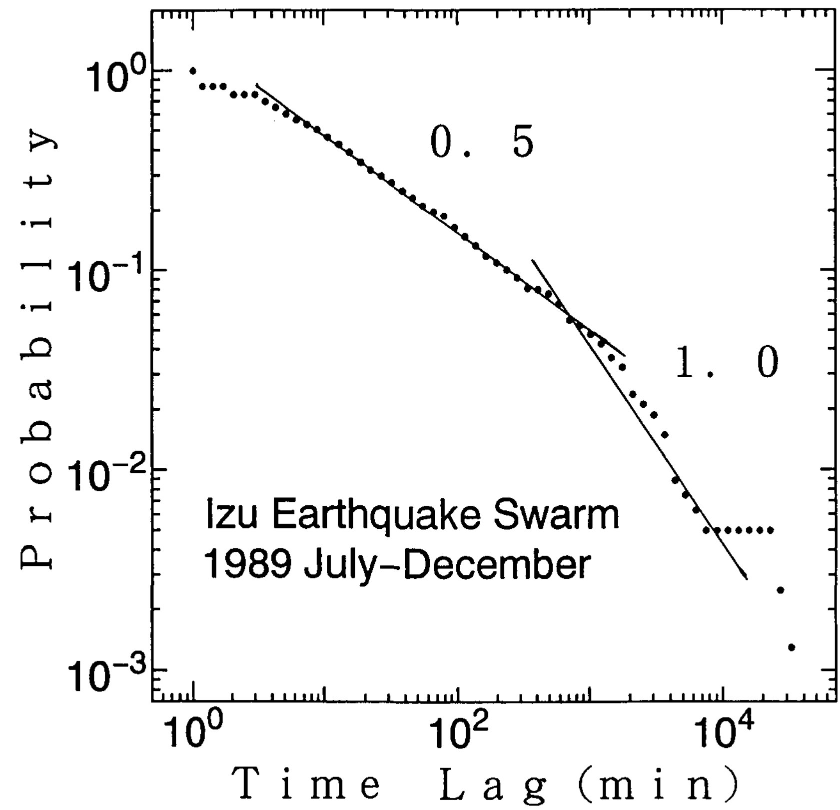


Fig. 5. Cumulative number of time-intervals of successive earthquakes in Izu peninsula, Japan for the period from July to December, 1989. Two straight lines indicate the power-law of cumulative number and time-interval in minutes.

earthquakes.

Another example in Fig. 5 is the probability of time-intervals for an earthquake swarm of more than 2000 local earthquakes in Izu peninsula, Japan from July to December of 1989. This also follows power-law distributions rather than the exponential function of Poisson distribution. Power coefficients of d_1 and d_2 in Fig. 5 are about 0.5 and a little smaller than 1.0. This swarm activity has a b-value for the G-R relation of about 0.82, while the b-value of the earthquake activity in the vicinity of Japan shown in Fig. 4 is about 0.85. The convex function, slower decay than that of a Poisson but not as slow as a simple power law in Figs. 4 and 5, will be further investigated later in this section.

3.2 *Modeling*

No previous theory has resolved a connection between Poisson and power-law distributions for earthquake occurrences. We can show the relation between the autocorrelation function and the time-interval distribution of the random processes. Therefore, we are able to consider jointly the long-tail behavior of the autocorrelation by Ogata and Abe (1991) and of the power-law of time-intervals. Note that the power coefficient of $2\text{-}2H$ with $H=0.6$ yields a power coefficient d of 0.8, which is consistent with the power coefficient of 0.5 and 1.0 for the earthquake swarm in Fig. 5.

The stochastic scaling in the previous section provides a clear method for reducing the family of exponential functions to a power-law function. Suppose that the probability density of successive events within a time-interval from τ to $\tau+d\tau$ is $P_0(\tau)d\tau$, then the Poisson process provides

$$P_0(\tau)=\lambda_0 \exp(-\lambda_0\tau) \ (\tau\geq 0), \tag{11}$$

where λ_0 is an average number of events within a unit of time. Let us consider that (11) represents the probability density of one series of earthquake occurrences classified by the minimum magnitude range. Another series of earthquakes classified by a slightly larger magnitude-range than that in (11) can be considered. The relative probability density of this series of earthquake activity can be expressed similarly to the stochastic scaling of (5) as

$$f_s: \frac{\alpha_s}{\beta_s} P_i(\beta_s\tau)=P_{i+1}(\tau) \ (i=0, 1, 2, \dots, M), \tag{12}$$

where α_s and β_s are positive scaling parameters. This scaling is also stochastic, because the functions which scale are probability functions. Within the time-interval from τ to $\tau+d\tau$, the i-th series of earthquakes has the probability density of P_i defined relatively to P_0 of the 0-th series in (12). Consequently, when the total event number of the 0-th series in (11) is considered, the relative number of the i-th series of activity can be obtained from (12). For sufficient number of events, there is no preference for any series of events. Sometime, an event of the $(i+1)$-st series will occur, and at other time one of the $(i+2)$-nd series will occur. Since each series of the activity is independent, the chance of an event does not depend on the probability for the other series of earthquakes. Therefore, the chance of an event from the cluster of these series is proportional to

$$P_s(\tau)=\sum_{i=0}^{M} P_i(\tau), \tag{13}$$

where $P_i(\tau)$ can be obtained formally from (11) and (12) as

$$P_i(\tau)=\lambda_0\left(\frac{\alpha_s}{\beta_s}\right)^i \exp(-\lambda_0\beta_s^i\,\tau)\ (\tau\geq 0). \tag{14}$$

The above equation of (13) indicates, in other words, the summed-up number of all the events expected within the time interval. Since the total number of events classified into each series decreases as the scaling proceeds, we have a constraint after integrating $P_i(\tau)$ in (14) with respect to τ from 0 to ∞ as

$$\frac{\alpha_s}{\beta_s^{\,2}}\leq 1. \tag{15}$$

Clearly, the above ratio relates to the scaling parameters α_d and β_d in (5) and to the empirical b-value in (2) as

$$\frac{\alpha_s}{\beta_s^{\,2}}=\frac{\alpha_d}{\beta_d}=10^{-b\varDelta M}, \tag{16}$$

where $\varDelta M$ is the magnitude interval. This relation provides an understanding of the physical meaning of the empirical b-value in terms of the stochastic scaling in (12).

An asymptotic form of (13) can be derived similarly by the steepest descent method

$$P_s(\tau)\simeq A_s|\lambda_0\,\tau|^{-\xi_s+1}, \tag{17}$$

$$A_s=\sqrt{\frac{\pi}{2\ln(\alpha_s/\beta_s)\ln\beta_s}}\,\lambda_0\exp([\,\xi_s-1]\,[\ln(\xi_s-1)-1]), \tag{18}$$

where ξ_s is a scaling dimension of the present stochastic scaling in (12), is defined as

$$\xi_s = \frac{\ln \alpha_s}{\ln \beta_s}. \tag{19}$$

Inequality of (15) restricts this scaling dimension $\xi_s \leq 2$. Although the asymptotic form of (17) is valid for the case of $\xi_s > 1$, the power-law can be also extended for the case of $1 > \xi_s > 0$ (Koyama and Hara, 1992). Consequently, although each series of earthquake activity, classified by a particular magnitude-range, is characterized by a Poisson process, the cluster of many of these series exhibits a power-law distribution of the autocorrelation and time-interval. Therefore, the power coefficient of d_1 and d_2 empirically obtained from observations is formally understood in relation to the scaling dimension ξ_s of the complex earthquake activity. Since the empirical d is obtained from cumulative plot and P_s in (17) is density, the relation is

$$d = 2 - \xi_s. \tag{20}$$

Once we determine the power coefficient of the time-interval distribution of an earthquake activity, the coefficient indicates the scaling property through (20). We conclude here that the fundamental property of complex earthquake activity can be simply described by the stochastic scaling. This result has been derived in a general manner, so we expect that this theory describes not only the long-tail behavior of the complicated earthquake activity but also the complex system of natural phenomena.

3.3 Log-normal and scaled-sum distributions

Foreshock activity of an earthquake is very important to predict the following mainshock. Since the number of foreshocks is usually very small, the difference between foreshock activity and earthquake swarm is not obviously understood. In case of the Haicheng earthquake of 1975 in China, more than 300 foreshocks were detected. The observation enabled to issue the earthquake warning and minimized the disaster. A b-value of foreshocks has been measured at about 0.56 (Wu, 1990; reproduced in Fig. 6 (a)), whereas the value of aftershocks in Fig. 6 (a) is measured at about 0.9.

Figure 6 (b) shows the time-interval distribution of successive foreshocks and aftershocks of the Haicheng earthquake. The power-law with a coefficient of about 0.8 fits the distribution of foreshocks but the exponential distribution would not. The power coefficient of aftershocks is not well determined, since the decay rates of medium and large time-intervals are very small compared with the exponential but not as small as the decay of a power-law. A common analysis for such asymmetric data is to apply a log-normal distribution, since

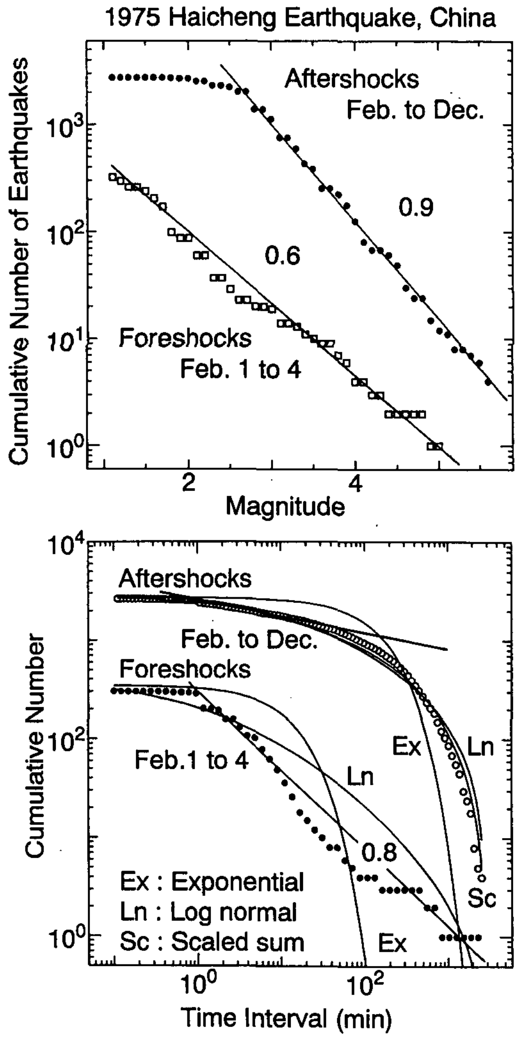


Fig. 6. (a) : Cumulative number of foreshocks and aftershocks of the 1975 Haicheng earthquake in China as a function of magnitude. Empirical estimate of b-value is about 0.6 for foreshocks and about 0.9 for aftershocks. The original data was provided by the State Seismological Bureau of China.

(b) : Cumulative number of foreshocks and aftershocks of the 1975 Haicheng earthquake in China as a function of successive time-interval. Solid line indicates a power-law with coefficient of about 0.8 for foreshocks. Solid curves labeled Ex indicate exponential distribution and those labeled LN indicate log-normal distribution for foreshocks and aftershocks. Scaled sum probability (labeled Sc) by the stochastic scaling is also derived from the aftershock data.

the distribution has both a long-tail and statistical moments.

In addition to foreshock/aftershock activities, Fig. 7 illustrates the time-interval distribution of the Matsumae earthquake swarm in Hokkaido, Japan. The swarm activity started on October, 1995 and the hypocenters of more than 4500 has been determined until early July of 1996 by the Research Center for Earthquake Prediction of Hokkaido University. The maximum magnitude registered 4.6 on November 23, 1995. The plot in Fig. 7 does not favor the exponential nor power-law distributions. The log-normal distribution may also explain the asymmetric distribution of plots in Fig. 7. This result encour-

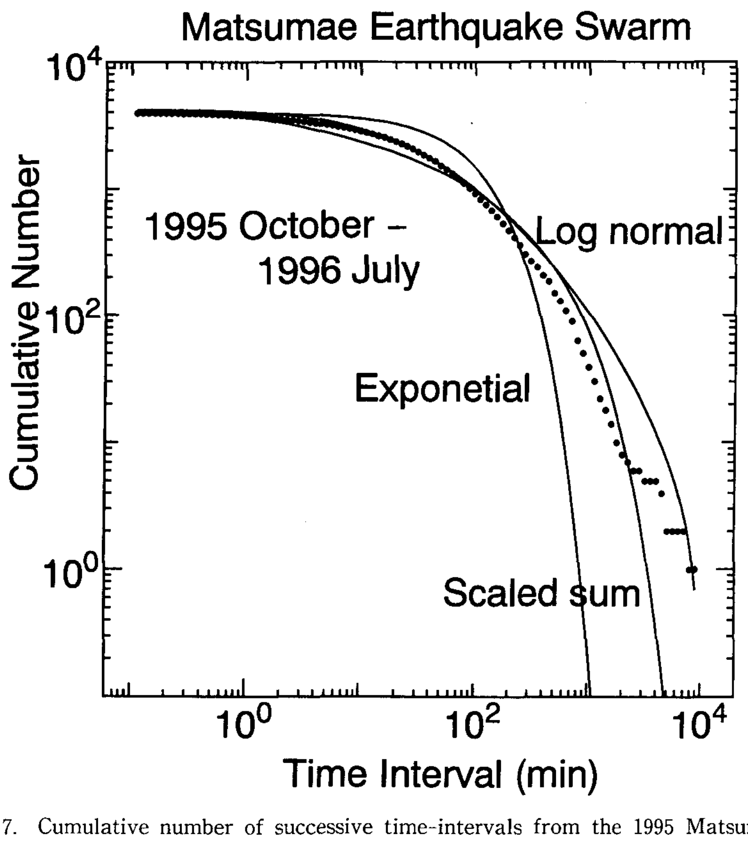


Fig. 7. Cumulative number of successive time-intervals from the 1995 Matsumae earthquake swarm, Hokkaido, Japan. All the events (more than 4500) from October of 1995 to early July of 1996 are analyzed, whose hypocenters have been determined by Research Center for Earthquake Prediction, Hokkaido University. Exponential and log-normal distributions as well as scaled-sum distribution by the stochastic scaling are shown in the figure.

ages us to reconsider two power-law distributions in Fig. 5.

Log-normal distributions have been observed in many diverse fields, and the time-interval distribution of successive aftershocks with variety of sizes and earthquake swarm seems to be another example as shown in Fig. 6 (b) and in Figs. 5 and 7. We reconsider this asymmetric distribution from the viewpoint of the stochastic scaling.

If ln x has a normal distribution then the variable x has the distribution

$$g\left(\frac{x}{\bar{x}}\right)\frac{dx}{\bar{x}}=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{[\ln(x/\bar{x})]^2}{2\sigma^2}\right)\frac{dx/\bar{x}}{x/\bar{x}}. \tag{21}$$

When the variance σ^2 is large and/or the variable x is close to the average $\bar{x}$, $g(x)$ mimics a $1/x$ distribution (Montroll and Shlesinger, 1983). The larger σ, the more orders of magnitude the mimicking persists. The $1/x$ distribution can be derived as an asymptote (17) when $\alpha_s \to \beta_s^2$. Therefore, the log-normal distribution can be approximated by our stochastic scaling of (12). Provided that $\alpha_s = \beta_s^2$, β_s (<1.0) is the only parameter. Scaled-sum distribution by the stochastic scaling has been calculated taking different values of variance σ^2 in (21), and two probability distributions were compared each other. The numerical result confirms the consistency between the log-normal and scaled-sum distributions for the sufficient amount of probability. The rest of few percent is quite different because of the difference in functional forms. This statement can be found in Fig. 6 (b) and Fig. 7.

The convex distribution of regional earthquakes in the vicinity of Japan (Fig. 4) and that of the Haicheng aftershocks (Fig. 6) is so understood by the stochastic scaling of the cluster of earthquake activities. The result is shown in Fig. 6 (b) and in Fig. 7, resolving the consistency of log-normal and scaled-sum distributions fitted to the data. Clearly found is that the two distributions are consistent each other explaining the observation in the range more than two order of magnitude.

The log-normal distribution is purely statistical, meanwhile the scaled-sum distribution is more physical. The fundamental property underlying the distribution can be physically understood by the scaling relation. Therefore, the present approach would give an insight into the phenomena that have been characterized by the log-normal distribution.

4. Maximum entropy of earthquake activity

Empirical values of p, b, power-coefficient d's and H have been numerically

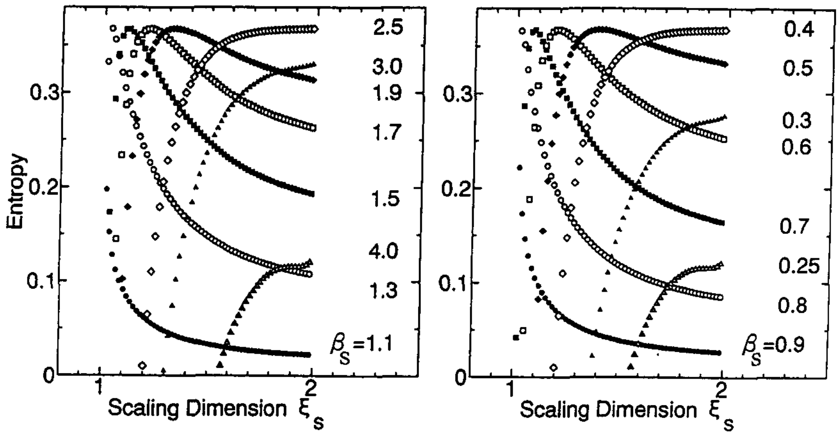


Fig. 8. Entropy of the complex earthquake activity in terms of scaling dimension. The scaling parameter β_s is also shown.

evaluated for natural earthquakes, and (16) relates the scaling parameters to b-value. Therefore, all the scaling parameters α_d, β_d, α_s and β_s can be determined for each earthquake activity. The b-value is usually about 0.8 to 1.0 and the coefficient p is about 1 to 2 for most of natural earthquake activities irrespective of the size and energy. Here we investigate how these values are preferentially determined in nature.

Since the asymptotic form in (17) is for $\lambda_0\tau \gg 1$, the time interval of $0 \leq \tau < \tau_0 (=1/\lambda_0)$ is not important for describing the complexity of the earthquake activity. This time interval is dependent mostly on the initial condition, when $\alpha_s < \beta_s < 1$. In the time interval of $(\lambda_0 \beta_s^M)^{-1} < \tau$, there is low probability for events, because of the small number of earthquake nuclei remaining. This introduces another time constant of $\tau_M = (\lambda_0 \beta_s^M)^{-1}$. These two characteristic times, τ_0 and τ_M specify the scaling region for the complex earthquake activity. When $\alpha_s > \beta_s > 1$, the time interval $\tau_M < \tau < \tau_0$ is the scaling region.

All the chances (probability) within this scaling region of time are calculated as

$$P_T = \int_{\tau_0}^{\tau_M} P_s(\tau) d\tau \Big/ \int_0^{\infty} P_s(\tau) d\tau. \tag{22}$$

The entropy of this cluster of the earthquake activity is then defined as

$$E_T = - P_T \ln P_T. \tag{23}$$

Under the condition in (15), we can calculate E_T for each pair of α_s and β_s.

Figure 8 shows the entropy (23) in terms of the scaling dimension ξ_s where the scaling parameters are $1<\beta_s<\alpha_s$ in the left and for $1>\beta_s>\alpha_s$ in the right. There are three cases in Fig. 8 for which the entropy (23) becomes maximum: The first case is for $\xi_s \to 1$ and β_s tends to 1.0. The second case for ξ_s about 1.1 to 1.4 and the third is for $\xi_s \to 2$.

The first case of the maximum entropy can be readily understood. Since the scaling parameter α_s also tends to 1.0 in this case, this is just the case without the scaling. All the series comprising the cluster of activity are the same. Therefore, each series in this case occurs at random with an equal probability, and obviously the entropy of such activity is maximum. Meanwhile, the third case indicates equal probability of series of activity, even though the component series are scaled by the scaling parameters of α_s and β_s. Since the latter case manifests $\alpha_s/\beta_s^2 \to 1$ instead of the inequality in (15), the probability and/or the total number of events of each cluster is the same. It is also reasonable that this case corresponds to the maximum entropy. In the second case, the complexity of the earthquake activity gives rise to the maximum entropy.

Since an earthquake swarm occurs in a limited source area, the magnitude range is usually small. Therefore, the summation of the series of earthquake activities covers a narrow range of magnitudes. This is an extreme case without scaling, which yields a power-coefficient d of about 1.0 and $\xi_s \to 1$. This is the first case for the maximum entropy of complex earthquake activity.

For a typical value of $b=0.8$ with $\Delta M=0.2$ and power-coefficient $d=0.8$ for the world-wide seismicity, the scaling dimension $\xi_s(=2-d)$ is calculated to be 1.2. In this case we observe α_s of about 1.74 and β_s of about 1.58 from (16) and (19). These values are almost the same as those for the earthquake swarm in Fig. 5, where d of $2\text{-}\xi_s$ is about 0.8. In contrast to these observations, we will study the theoretical background. Suppose that β_s is 1.6, the theoretical scaling dimension for the second case of the maximum entropy is about 1.2 in Fig. 8. Away from the above example, provided that $b=1.0$, $\Delta M=0.1$ and $d=0.8$, the case leads to $\alpha_s=1.41$, $\beta_s=1.33$ and $\xi_s=1.2$. The second maximum entropy for $\beta_s=1.33$ predicts ξ_s of a little less than 1.1 in Fig. 8. This indicates that the world-wide seismicity studied by Ogata and Abe (1991) and the earthquake swarm in Fig. 5 are well described by the maximum entropy complex activity.

Although the earthquake activity is complex, it is typically represented by the above b- and d-values which are expressed by the physical scaling parameters satisfying the condition of the maximum entropy. Note that the characteristic time of larger earthquakes is shorter than that of smaller earthquakes, since $\beta_s>1$ in this analysis. This may seem contradictory to our intuition, but this suggests that large aftershocks seldom occur and most of the aftershocks are very small long after the main shock.

For p of 1.1 and b of 0.8, for example, the fractal dimension of ξ_d is calculated to be 1.1. In this case α_d and β_d are about 0.017 and 0.025, respectively. For p of about 1.05, slightly larger than 1.0, we observe α_d of 0.0004 and β_d of about 0.0006. The scaling parameter β_d becomes smaller as $p \rightarrow 1$. Since β_d measures the time unit, this gives an extremely long time-constant for decaying aftershocks. From the relationship of $p = \xi_d$ with (16), we can show that scaling dimension ξ_s tends to 2 as $p \rightarrow 1$. Therefore, the long-tail behavior of aftershock activity corresponds to the third case of the maximum entropy. The stochastic scaling applied for the log-normal distribution also predicts that ξ_s tends to 2. So that, the aftershocks of the 1891 Nobi earthquake and the 1975 Haicheng earthquake are also classified to the third case of the maximum entropy.

Foreshock activities with a b-value as small as 0.5 have been found. The second case of the maximum entropy in Fig. 8 with $b=0.6$ and $d=0.9$ predicts ξ_s of about 1.1 and β_s of about 1.4. When we compare these theoretical values with the observations, the empirical values of b and d in Fig. 6 are consistent with the theoretical prediction. This is very important for understanding the essential property of foreshocks to forecast following major earthquakes.

5. Non-linear scaling law of earthquake activity

We have learned that the autocorrelation of earthquakes in and near Japan is represented by a power-law function which is identical to that of the worldwide seismicity (Ogata and Abe, 1991). We also showed the power-law distribution of time intervals of the earthquake swarm in Izu, Japan and of major earthquakes in and near Japan. Some are consistent with log-normal or scaled-sum distributions which are included within the natural extension of the power law. Power-law distributions of the local earthquake activity do not necessarily require a power-law distribution for the global earthquake activity. G-R relations for local earthquakes, regional earthquakes and global earthquakes have been obtained and they suggest the power law in variety of earthquake source sizes. This is another hierarchy represented by the power law, where the component subsets are also characterized by power laws. These observations indicate that there is some universal nature which connects power-law distributions of local (subset) activities with that of the global (full set) activity. Hara and Koyama (1992) investigated the scaling relationship between *local* and *global* complex systems in a general manner. We slightly modify the original theory (Hara et al., 1997) and apply it to the complex earthquake activity.

Suppose that there is a cluster of earthquake activities represented by a power-law autocorrelation as in Ogata and Abe (1991),

$$C_j(\tau) \propto \tau^{-\eta_j}. \tag{24}$$

For a global standpoint, there exist ω_j of such clusters characterized by the same value of η_j. Because the global activity is a summation of many different clusters and because clusters are considered to be independent, the global nature is expressed by the sum of (24) as

$$C(\tau) = \sum_{j=0}^{M} w_j C_j(\tau). \tag{25}$$

The above summation means in other words that there is no interaction among different clusters occurring in different places. This contrasts to independent series of activities within one cluster described by the stochastic scaling previously. The functional form of $C(\tau)$ must be

$$C(\tau) \propto \tau^{-q}, \tag{26}$$

in order to explain the power-law autocorrelation of the global earthquake activity, where q is a constant or may weakly depend on τ. This is what we understand from empirical power-law distributions for local and global earthquake activities, where η_j and q can be determined empirically.

In order to evaluate (25) and to clarify the relationship in terms of w_j, the clusters are re-labeled in increasing order of η_j as $0 \leq \eta_1 < \eta_2 < \dots < \eta_M$. Taking a new variable x of $\eta = \eta_M\, x (0 \leq x \leq 1)$, a formal representation of (24) and (26) is given as

$$\int_0^1 \tau^{-\eta_M x} w(x) dx \simeq \Psi \tau^{-q}, \tag{27}$$

where Ψ indicates the level of the global earthquake activity. Suppose that the weighting function $w(x)$ is scaled as

$$w(x) = \alpha\beta w(\beta x). \tag{28}$$

Since (27) is rewritten taking $\beta = \eta_M \ln \tau$,

$$\int_0^1 \exp(-\eta_M x \ln \tau) w(x) dx = \alpha \int_0^1 \exp(-\eta_M x \ln\tau) w(\eta_M x \ln\tau) d(\eta_M x \ln\tau) = \Psi \tau^{-q}, \tag{29}$$

we obtain (28) as

$$w(x) = \tau^{-q} \eta_M \ln\tau \; w(\eta_M x \ln\tau). \tag{30}$$

This equation represents the scaling relation of $w(x)$, and it leads to

$$w(x) \propto x^{\kappa(\tau)}. \tag{31}$$

The integrand of the left-hand side of (27) has a peak value when $x = \kappa(\tau)/\eta_M \ln \tau$ in the range of $0 \leq x < 1$. The integrand decreases monotonously when $x > \kappa(\tau)/\eta_M \ln \tau$. It is negligibly small when $x > 1$, so the integral in (27) can be extended approximately to infinity without losing generality. Then, (29) is

$$\int_0^\infty \exp(-\eta_M x \ln\tau)\, x^{\kappa(\tau)} dx = \tau^{-q} \int_0^\infty \exp(-X)\, w(X)\, dX, \tag{32}$$

where

$$X = \eta_M x \ln\tau, \tag{33}$$

$$\kappa(\tau) = \frac{q \ln\tau}{\ln|\eta_M \ln\tau|} - 1. \tag{34}$$

The last integral of the right-hand side of (32) is Γ function and it changes little, from about 0.8 to 2 with respect to the change of $\kappa(\tau)$ from -0.5 to 2. Therefore the non-linear weighting function of $w(x) \propto x^{\kappa(\tau)}$ provides the theoretical connection between the power laws of the local activity and the global activity.

Figure 9 shows $\kappa(\tau)$ as a function of τ. Here η_M is assumed to be 2, because η is equivalent to $2H-2$ with $0 < H \leq 1$. The power constant q is assumed to be 0.6, 0.8 and 1.0 so as to fit the actual observation. $\kappa(\tau)$ is a slowly-increasing function of τ. For $q = 0.8$ (the case in Ogata and Abe (1991)) for example, $\kappa(\tau)$ increases from about 0.7 to about 1.5 for a large change in τ from 100 to 10000. In a large lag-range, $\ln \tau$ of about 10 for example, $C(\tau)$ is composed of $0.689\tau^{-1.6}$, $0.216\tau^{-0.8}$ and $0.068\tau^{-0.4}$, respectively for $x = 0.8$, 0.4 and 0.2. The density of faster decay of $\tau^{-0.6}$ is ten times larger than that of the slower decay of $\tau^{-0.4}$. While in a small lag-range, τ of about 2 for example, $C(\tau)$ of $0.97\tau^{-0.6}$, $0.87\tau^{-0.8}$ and $0.78\tau^{-0.4}$. The density difference in this case is less than a factor of 1.3. These calculations suggest that the larger the lag is, the faster the autocovariance decays. This is the explanation for the convex plots which we have seen in Figs. 4, 5, 6 and 7 from the view point of the scaling of complex system.

The stochastic scaling in (5) or (12) describes the hierarchy of many series of local earthquake occurrences, and it is described by a linear mapping function. Therefore, such stochastic scaling applies to the local earthquake activity. However, the scaling of many of these clusters cannot be expressed by a simple linear function. Although we could not specify the mapping function,

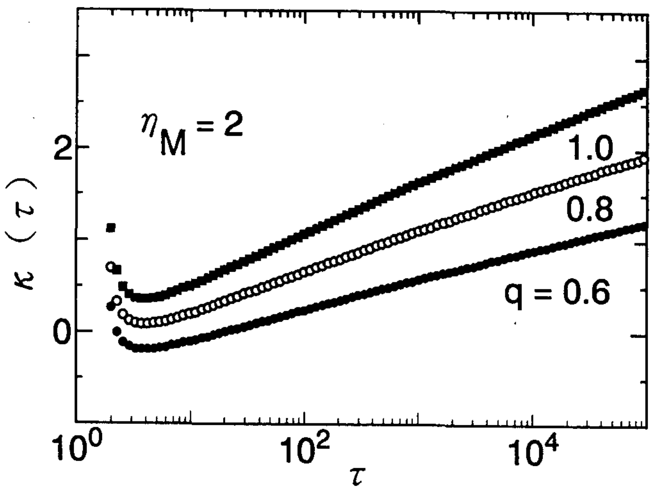


Fig. 9. Power coefficient $\kappa(\tau)$ of the non-linear scaling law in terms of time-lag τ.

the scaling involved with the weighting function of $w(x)$ is highly non-linear. This non-linear scaling applies to the global earthquake activity. Consequently, the stochastic scaling and the non-linear scaling fully describe the hierarchy of the complex earthquake activity on both local and global scales.

This non-linear scaling is different from the simple hierarchy of the complex system. The family of random phenomena characterized by exponential functions leads to the long-tail or slow dynamics represented by power-law functions. For this we have studied the stochastic scaling. The family of slow dynamics by power-laws is expected to be super-slow dynamics characterized by $(\ln t)^{-\zeta}$ (Hara et al., 1997). This can be understood as following relationships:

$$\Sigma\left(\frac{a}{b}\right)^j \exp(-b^j \bar{\gamma} t) \to t^{-\kappa},$$

$$\Sigma\left(\frac{a'}{b'}\right)^j \exp(-b'^j \chi \ln t) \to (\ln t)^{-\zeta},$$

because of $t^{-\kappa} = \exp(-\chi \ln t)$. Therefore, super-slow dynamics constitutes the hierarchy of complex systems, while the non-linear scaling law characterizes the composite system for the cluster of many complex systems.

6. Discussion and conclusions

Various time sequences of the complex earthquake activity have been studied by introducing a simple mathematical model. The basic principle is to apply the stochastic scaling to a cluster of many series of earthquakes classified by specific magnitude (energy) ranges. Each series is modeled by a stochastic process of random earthquake occurrences. The stochastic scaling describes the mutual relation among statistical properties of the series. Therefore, the complexity of the earthquake activity is represented by scaling parameters and the total number (or total energy) of the activity. Thus the physical basis of the present model is quite simple and far-reaching. Since the stochastic scaling is considered in the fundamental equation without any externally-imposed stochasticity or heterogeneity, we are able to evaluate the total energy and the entropy of the complex earthquake activity. Without this basic approach, we cannot determine how the scaling parameters are preferentially determined in nature.

Since the earthquake swarm occurs in a limited source area, the magnitude range of earthquakes is usually very narrow. Therefore, the stochastic size-effect is not essential to explain the temporal variation of the earthquake swarm activity. This reduces the temporal variation to the production and reduction of latent sources of earthquakes in the crust and to the exponential decay of the activity. This is important in understanding the complex earthquake activity, because the entropy of this type of swarm activity is maximum.

Magnitudes of aftershocks of large earthquakes cover a wide range. It is reasonable to apply the asymptotic form (7) in such size-dependent random phenomena. Consequently, the empirical power-law decay of aftershocks is understood in terms of the complexity of the earthquake activity. The stochastic scaling here predicts Omori-Utsu empirical constant to be in the range of $1<p<2$. For $p<1$ the stochastic scaling does not apply. The long-tail of aftershock activity is observed when p tends 1.0 and the activity is also characterized by the maximum entropy.

The present study differs from previous analyses dealing with non-linear dynamical systems to simulate the earthquake occurrence in the manner above. Even for a small degrees of freedom, non-linear dynamical systems exhibit chaotic behavior showing the complicated pattern of earthquake-like occurrences. Previous models predict a size-number distribution of random phenomena similar to what is seen in the earthquake activity (e.g., Bak and Tang, 1989; Carlson, 1991; Rundle, 1989). However, the random phenomena from these model simulations is not always applicable to the earthquake phenomenon. For example, the slip velocity on a fault plane can be evaluated as a random variable by the numerical simulations of the previous models of the self-organized criticality. Slip velocity on natural faults is almost constant and does not change by as much as a factor of 2. This is an important constraint for understanding the kinematics of heterogeneous earthquake faulting processes (Koyama, 1994).

Obviously, the basic features that are observed in the previous models should be representative of a wide class of physical models. These provide clues for understanding the generation mechanism of the earthquake phenomenon, one of which has been studied by Ouchi (1993). Nevertheless, further insight is needed before comparing directly the results obtained from the models with those of natural earthquakes.

The strength and the number of earthquakes are statistically scaled in this study. We have not discussed on the physical origin of the scalings, however the scaling property can be found in the diversity of the complex systems in nature. The stochastic scaling here is one such scaling representing a fundamental aspect of the complex system. As a consequence of this scaling, we understand earthquake swarms as rupture processes of fault patches without the scaling or with a characteristic short-wavelength. This is consistent with the simulation showing that a highly inhomogeneous stress field with short-wavelengths increases the frequency of earthquake occurrences. Since the stochastic scaling performs a coarse-graining of statistical properties of earthquake occurrences, an earthquake sequence of foreshocks, mainshocks and aftershocks is the result of random occurrences of events, which constitutes the local and global seismicities.

The long-tail behavior of the local, regional and global seismicities is the evidence of the hierarchy structure of earthquake occurrences. The essential property of the hierarchy is described by the non-linear scaling law. The non-linear scaling law explains the variation in the power coefficients in Fig. 5. The present theory has been applied to the autocovariance function, and the theory can be also applicable to Gutenberg and Richter's relations for small earthquakes and for great earthquakes in the world. Therefore, the b-value variation of natural earthquakes, which has been pointed out recently by Pacheco et al. (1992) can be similarly understood.

Acknowledgement

This study has been developed while one of the authors (J. K.) was an exchange scientist of the Peking University. He would like to thank National Education Committee of People's Republic of China and Japan Society for the Promotion of Science. He would like to thank Prof. Zang Shaoxian of the Peking University for his warm hospitality. Dr. Alan Linde, Department of Terrestrial Magnetism, Carnegie Institution, kindly read the manuscript and suggested many improvements to revise it. This study is partly supported by Grant No. 08640520 in Aid for Scientific Research, the Ministry of Education of Japan.

References

Aki, K., 1956. Up-to-date earthquake statistics (in Japanese), Zisin Ser. 2, **8**, 205-228.

Bak, P. and C. Tang, 1989. Earthquakes as self-organized critical phenomenon, J. Geophys. Res., **94**, 635-637.

Brown, S. and C.H. Sholtz, 1985. Broad bandwidth study of the topography of natural rock surfaces, J. Geophys. Res., **90**, 12575-12582.

Burridge, R. and L. Knopoff, 1967. Model and theoretical seismicity, Bull. Seism. Soc. Am., **57**, 341-371.

Carlson, J.M., 1991. Time intervals between characteristic earthquakes and correlations with smaller events: An analysis based on a mechanical model of a fault, J. Geophys. Res., **96**, 4255-4267.

Cheng, Y.T. and L. Knopoff, 1987. Simulation of earthquake sequences, Geophys. J. Roy. astr. Soc., **91**, 693-709.

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

Feder, J., 1988. *Fractals*, Plenum Press, New York, 1-283.

Gu J.C., X.B. Xie and L. Zhao, 1979. On time distribution of large aftershocks of the sequence of a major earthquake and preliminary theoretical explanation (in Chinese), Acta Geophys. Sinica, **22**, 32-45.

Gutenberg, B. and C.F. Richter, 1954. *Seismicity of the Earth*, Princeton University Press, Princeton, 1-310.

Hara, H. and J. Koyama, 1992. Activation of complex system and its autocorrelation (in Japanese), Statistical Mathematics, **40**, 217-226.

Hara, H., S.S. Lee, J. Koyama and S. Fujita, 1997. Scaled Langevin equation for complex system: New linear scaling relation for weight factor, Chaos, Solitons & Faractals, 8, 1685-1691.

Koyama, J., 1994. General description of the complex faulting process and some empirical relations in seismology, J. Phys. Earth, **42**, 103-148.

Koyama, J., 1997. *The complex faulting process of earthquakes*, Kluwer Academic Publishers, Dordrecht, 1-194.

Koyama, J. and D.Y. Feng, 1995. *Advance in Mathematical Seismology*, Seismological Press, Beijing, 1-193.

Koyama, J. and H. Hara, 1992. Scaled Langevin equation to describe $1/f^a$ spectrum, Phys.

Rev. A, **46**, 1844-1849.

Koyama, J. and H. Hara, 1993. Fractional Brownian motions described by scaled Langevin equation, Chaos, Solitons & Fractals, **3**, 467-480.

Mogi, K., 1991. Irreversible crustal activity in Izu peninsula (in Japanese), Earth Monthly, **13**, 55-62.

Montroll, E.W. and M.F. Shlesinger, 1983. Maximum entropy formalism, fractals, scaling phenomena, and $1/f$ noise: A tale of tails, J. Stat. Phys., **32**, 209-230.

Ogata, Y. and K. Abe, 1991. Some statistical features of the long-term variation of the global and regional seismic activity, Intern. Stat. Rev., **59**, 139-161.

Ohtsuka, M., 1972. A chain-reaction type source model as a tool to interpret the magnitude-frequency relation of earthquakes, J. Phys. Earth, **20**, 35-45.

Ouchi, T., 1993. Population dynamics of earthquakes and mathematical modeling, PAGEOPH, **140**, 15-28.

Pacheco, J.F., C.H. Scholz and L.R. Sykes, 1992. Changes in frequency-size relationship from small to large earthquakes, Nature, **355**, 71-73.

Ruff, L.J., 1992. Asperity distribution and large earthquake occurrences in subduction zones, Tectonophysics, **211**, 61-84.

Rundle, J.B., 1989. Derivation of the complete Gutenberg-Richter magnitude-frequency relation using the principle of scale invariance, J. Geophys. Res., **94**, 12337-12342.

Tsukuda, T., 1993. Time sequence of the Nishi Omote jima earthquake swarm (in Japanese), Earth Monthly, **15**, 218-224.

Turcotte, D.L., 1997. *Fractals and chaos in geology and geophysics*, Cambridge University Press, Cambridge, 1-398.

Utsu, T., 1969. Aftershocks and earthquake statistics (I), J. Fac. Sci., Hokkaido University, **3**, 129-195.

Utsu, T., Y. Ogata and R.S. Matsuura, 1995. The centenary of the Omori formula for a decay law of aftershock activity, J. Phys. Earth, **43**, 1-33.

Vera-Jones, D., 1976. A branching model for crack propagation, PAGEOPH, **114**, 711-726.

Wang, J.H., 1994. On the correlation of observed Gutenberg-Richter's b value and Omori's p value for aftershocks, Bull. Seism. Soc. Am., **84**, 2008-2011.

Wu, K.T., 1990. *An Introduction to Earthquake Sequences* (in Chinese), Peking University Press, Beijing, 1-250.

Yamashita, T. and L. Knopoff, 1987. Models of aftershock occurrence, Geophys. J. Roy. astr. Soc., **91**, 13-26.

Yamashita, T. and L. Knopoff, 1992. Model for intermediate-term precursory clustering of earthquakes, J. Geophys. Res., **97**, 19873-19879.