The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific
and Cultural Organization

International Atomic
Energy Agency

SMR.1676 - 29

8th Workshop on Non-Linear Dynamics and Earthquake Prediction

3 - 15 October, 2005

Unified Laws in Seismicity: Existence and Optimal Spatial Scaling. Part 1

Interevent Time Distribution in Seismicity: A Theoretical Approach

George Molchan
Russian Academy of Sciences
International Inst. of Earthquake Prediction Theory and
Matematical Geophysics
Warshavskoye Sh. 79. Kor2
117556 Moscow
Russia

Pure appl. geophys. 162 (2005) 1135–1150
0033 – 4553/05/071135 – 16
DOI 10.1007/s00024-004-2664-5

Pure and Applied Geophysics

Interevent Time Distribution in Seismicity: A Theoretical Approach

G. MOLCHAN[1,2]

Abstract—This paper presents an analysis of the distribution of the time τ between two consecutive events in a stationary point process. The study is motivated by the discovery of unified scaling laws for τ for the case of seismic events. We demonstrate that these laws cannot exist simultaneously in a seismogenic area. Under very natural assumptions we show that if, after rescaling to ensure $E\tau = 1$, the interevent time has a universal distribution F, then F must be exponential. In other words, Corral's unified scaling law cannot exist in the whole range of time. In the framework of a general cluster model we discuss the parameterization of an empirical unified law and the physical meaning of the parameters involved.

Key words: Statistical seismology, point processes, self-similarity.

1. Introduction

Recently BAK *et al.* (2002) and CORRAL (2003a,b) suggested a new scaling law for seismic events on the phase space location-time-magnitude. The Corral's version of this law looks as follows: The distribution density for time τ between two consecutive events of magnitude $m > m_c$ has the form

$$p_\tau(t) = \lambda f(\lambda t), \tag{1}$$

where λ is the rate of events with $m > m_c$ in a given area G, while f is a universal function that is independent of the choice of G and cutoff magnitude m_c. The relation (1) is astonishing, being tested (as it has been by CORRAL, 2003b) for a very wide range of m_c (between 2 and 7.5), for seismic regions G of very different linear size L (between 20 km and the size of the Earth), as well as for different catalogs, both regional and global ones, and different time periods.

The parameterization of f seems not yet to have settled down. According to CORRAL (2003b):

[1] International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Warshavskoe shosse, 79, k.2, 117556, Moscow
[2] The Abdus Salam International Centre for Theoretical Physics, SAND Group, Trieste, Italy (e-mail: molchan@mitp.ru)

$$f(x) = cx^{\gamma-1}\exp(-x/a) \tag{2}$$

in the region $x \geq 0.05$ with $\gamma = 0.74 \pm 0.05$ and $a = 1.23 \pm 0.15$. The pioneering work (BAK *et al.*, 2002) uses the parameterization (2) for the whole range of x with $\gamma = 0.1$ (see refined estimates in CORRAL, 2003a). This allows the behavior of $f(x)$ about zero to be interpreted in terms of the Omori law.

The subsequent discussion strives to answer the following questions.

What is the behavior of the distribution of τ near 0 and ∞ in stochastic models of seismicity? These parts of the distribution of τ usually dominate the log-log representation and therefore are important for understanding the unified law in the framework of the classical models. The next question is: what is the physical meaning of the parameters γ and a in (2)? Finally, assuming the form of f to be universal for τ, what should it be?

The answers to these questions are contained in Sections 2–4. All proofs of the main results are collected in the Appendix.

2. *A Poisson Cluster Model: The Asymptotics of $p_\tau(t)$*

Earthquakes frequently form anomalous clusters in spacetime. The largest event in a cluster is termed the *main* event. The events that occurred before and after the main event in a cluster are called fore- and aftershocks, respectively. It is assumed in a zero approximation that main events constitute a time-uniform Poisson process. That assumption is widely employed in seismic risk studies.

Aftershocks dominate clusters both as regards their number and duration. Their rate as a function of time is described by the Omori law:

$$n(t) = ct^{-p}, \quad t > t_0, \tag{3}$$

where t_0 is small. Relation (3) holds fairly well during the first few tens of days (up to a year) with the parameter $p = 0.7 - 1.4$ (UTSU *et al.*, 1994) At large times the value of p becomes greater, occasionally significantly so, making $n(t)$ decay in an exponential manner. Taken on the whole, background seismicity and spatial interaction do not allow reliable conclusions to be drawn for the Omori law at large times. Cases in which (3) holds during decades are unique (UTSU *et al.*, 1994).

Following the above description, we consider the following model for seismic events in time. The spatial and magnitude components of events are disregarded for simplicity of reasoning. Let $\{x_i\}$ be a homogeneous Poisson point process on a line with rate λ^*. It is an analogue of main events. Let $N_0(dt)$ be an inhomogeneous point process with rate $\delta(t) + \lambda_0(t)$. Here, δ is the delta function, while the presence of $\delta(t)$ means that the event $t = 0$ belongs to N_0. The notation $N_0(\Delta)$ defines the number of events N_0 in the interval Δ. We will assume that

$$\int \lambda_0(t)dt = \Lambda < \infty. \tag{4}$$

This requirement ensures that the total number of events in N_0 is a.s. bounded.

Consider an infinite series $N_0^{(i)}(dt)$, $i = 0, \pm 1, \pm 2, \ldots$ of independent samples of N_0. The theoretical process N is the sum

$$N(dt) = \sum_i N_0^{(i)}(dt - x_i).$$

The process $N_0^{(i)}$ that has been shifted by the amount x_i can be associated with the cluster of the main event x_i.

Our task is to describe the distribution of τ between two consecutive events in N. The distribution is uniquely specified, because the process N is stationary. It is also easy to see that the rate of N is

$$\lambda = \lambda^*(1 + \Lambda).$$

According to (DALEY and VERE-JONES, 2003),

$$P(\tau > t) = -\frac{d}{\lambda dt} P\{N(0,t) = 0\} \tag{5}$$

and

$$P(N(\Delta) = 0) = \exp\Big\{ -\lambda^* \int P(N_0(\Delta - x) > 0)dx\Big\}.$$

The first relation is a version of the Palm-Khinchin equation (3.4.9) in DALEY and VERE-JONES (2003) appropriate to general stationary point processes, while the second is based on the fact that the main events are Poissonian (see (6.3.15), p.181 ibid). Since the aftershocks make the bulk of a cluster, we shall assume in what follows that $\lambda_0(t) = 0$ for $t < 0$. Consequently,

$$P(N_0(\Delta) > 0) = \begin{cases} 1, & \text{if} \quad 0 \in \Delta \\ 0, & \text{if} \quad \Delta \subset (-\infty, 0). \end{cases}$$

Combining the above relations, we get

$$P(\tau > t) = \exp\{-\lambda^* \int_0^\infty P(N_0(u, t+u) > 0)du - \lambda^* t\} \times$$
$$[1 + \int_0^\infty P(N_0(t + du) > 0, N_0(u, u+t) = 0)]/(1+\Lambda). \tag{6}$$

We now describe the behavior of the distribution of τ near 0 and ∞.

Statement 1. (a) *If cluster duration has a finite mean, $\bar{\tau}_{cl}$, then*

$$P(\tau > t) = \exp(-\lambda^*(t + \bar{\tau}_{cl}))/(1 + \Lambda) \cdot (1 + o(1)), \quad t \to \infty.$$

(b) *Let* $\lambda_0(t) \sim ct^{-1-\theta}$, $t \to \infty$ where $0 < \theta < 1$. Then

$$P(\tau > t) = \exp(-\lambda^* t - O(t^{1-\theta}))/(1 + \Lambda), \quad t \to \infty. \tag{7}$$

In other words, one has

$$\lim_{t\to\infty} \ln P(\tau > t)/(\lambda t) = -\lambda^*/\lambda$$

for a Poisson sequence of main events in a broad class of cluster models. In terms of the parameterization (2), that means that

$$a = \lambda/\lambda^* = 1 + \Lambda.$$

With $a = 1.23$ (as in CORRAL, 2003b) the main events constitute $a^{-1} \simeq 81\%$ of the total number of events.

The following regularity conditions should be imposed on N_0 in order to be able to describe how the distribution density for τ behaves for small t:

$$P(N_0(u, u+t) > 0 | N_0\{u+t\} = 1) = o(1), t \to 0 \tag{8}$$

$$P(N_0(u, u+t) > 0 | N_0\{u\} = 1, N_0\{u+t\} = 1) = o(1), t \to 0, \tag{9}$$

where the notation | denotes conditional probability, and $N_0\{s\} = 1$ means that there is an event at the point s. We assume in addition that (8), (9) hold uniformly in $u \geq 0$.

That last requirement is no limitation for the case of seismic events, considering that the rate of cluster events and time relations between them seem to be rapidly decaying over time. The requirements (8), (9) themselves ensure that two very closely lying cluster events are not likely to contain another cluster event between them. However, the requirements (8), (9) are different from the following orderliness condition in DALEY and VERE-JONES 2003: $P\{N_0(u, u+t) > 1\} = o(1), \ t \to 0$ uniformly in u.

Statement 2. *If* (8), (9) *hold, the probability density for* τ *(provided it exists) has the following form as* $t \to 0$:

$$p_\tau(t) = \left[\lambda_0(t) + \int_0^\infty \lambda_0(u)\lambda_u(t)du + \lambda(1+\Lambda)\right]/(1+\Lambda) \cdot (1 + o(1)), \tag{10}$$

where $\lambda_u(t) = P(N_0(t+u, t+u+\delta) > 0 | N_0\{u\} = 1)/\delta$, $\delta \ll 1$, *is the conditional rate of* N_0 *after time u given a cluster event has occurred at that time. In particular, if* $\lambda_0(t) \uparrow \infty$ as $t \to 0$ *and, for some finite* $k > 0$,

$$\lambda_u(t) < k\lambda_0(t), \quad 0 < t < \varepsilon, \tag{11}$$

then

$$1 < p_\tau(t)/\lambda_0(t) < c, \quad t \to 0.$$

In other words, when (8), (9) hold, the distribution density of τ for small t is proportional to the rate of cluster events immediately after the main event. The statement is not obvious, since any interevent interval is not necessarily started by a main event.

3. Examples

Examples will now be discussed so as to understand to what extent the above assumptions are restrictive.

The trigger model. Historically, this was the first seismicity model to appear (see VERE-JONES, 1970). It assumes the cluster process N_0 to be Poissonian. The model has not found acceptance in seismicity statistics, because the likelihood of an observed sample in that model is technically difficult to use. This does not rule out that the model may be helpful, however.

Because increments in N_0 are independent, the requirements (8), (9) have the form

$$P(N_0(u, u+t) > 0) = 1 - \exp\left(-\int_u^{u+t} \lambda_0(x)dx \right) = o(1), \quad t \to 0.$$

If $\lambda_0(x)$ is a decreasing function, one has

$$\int_u^{u+t} \lambda_0(x)dx < \int_0^t \lambda_0(v)dv = o(1).$$

Consequently, the decrease of $\lambda_0(x)$ ensures that (8), (9) hold uniformly in u. The same property of $\lambda_0(x)$ also ensures (11):

$$\lambda_u(t) = \lambda_0(u+t) < \lambda_0(t).$$

We now refine the asymptotic form of $p_\tau(t)$ for small t.

Let $\lambda_0(x)$ be a smooth decreasing function and $\lambda_0(t) = ct^{-p}$, $t < 1$. Then

$$p_\tau(t) \simeq ct^{-p} + c_1 t^{-\alpha} + c_2, \quad t \to 0,$$

where $\alpha = 2p - 1$ for $p > 1/2$ and $\alpha = 0$ for $p \leq 1/2$.

This can be seen as follows. When $p > 1/2$, one has

$$I_t = \int_0^\infty \lambda_0(u)\lambda_u(t) = c^2 \int_0^1 u^{-p}(u+t)^{-p}du + \int_1^\infty \lambda_0(u)\lambda_0(u+t)du$$

$$= c^2 t^{1-2p} \int_0^\infty u^{-p}(1-u)^{-p}du + \text{const} \; + o(1), \quad t \to 0.$$

When $p < 1/2$, one has

$$I_t = \int_0^\infty \lambda_0^2(u)du + o(1).$$

The self-exciting model. A cluster in this model is generated by the following cascade process. The first event $t = 0$ is defined as the event of rank 0. It generates a Poisson process with rate $\pi_0(t)$; its events $\{t_i^{(1)}\}$ are ascribed rank 1. The procedure then becomes recursive: each event $\{t_i^{(r)}\}$ of rank $r = 1, 2, \ldots$ generates a Poisson process of its own which is independent of the previous ones and which has the rate $\pi_0(t - t_i^{(r)})$. The offspring of a rank r event are events of rank $r + 1$, the events of all ranks constituting the desired cluster N_0.

The process N with clusters as described above is known as the self-exciting model (HAWKES and ADAMOPOULOS, 1973) or the epidemic type model (OGATA *et al.*, 1986) The model is rather popular in statistical studies and forecasting of seismicity thanks to the fact that the predictable component of N has simple structure:

$$E(N(t+\delta) - N(t) > 0|\mathscr{A}_t) = \sum_{t_i < t} \pi_0(t - t_i) \cdot \delta + \lambda^* \delta, \quad \delta \ll 1,$$

where the t_i are events of $N(dt)$ and $\mathscr{A}_t = \{N(ds), s < t\}$ is the past of the process.

It is easy to see that the rate λ of the process N is bounded, if

$$\lambda_\pi = \int_0^\infty \pi_0(t) dt < 1;$$

also,

$$\Lambda = \lambda_\pi/(1 - \lambda_\pi) \quad \text{and} \quad \lambda = \lambda^*/(1 - \lambda_\pi).$$

Statement 3. (a) *The cluster rate function for the self-exciting model is*

$$\lambda_0(t) = \pi_0(t) + \pi_0 * \pi_0(t) + \pi_0 * \pi_0 * \pi_0(t) + \ldots, \quad t > 0, \tag{12}$$

*where * denotes the convolution.*

Let $\pi_0(t)$ be monotone near 0, *where $\pi_0(t) \sim At^{-p}$, $0 < p < 1$. Then*

$$\lambda_0(t)/\pi_0(t) \sim 1, \quad t \to 0.$$

Let $\pi_0(t)$ be monotone at ∞, where $\pi_0(t) \sim Bt^{-1-\theta}$, $0 < \theta < 1$. Then

$$\lambda_0(t)/\pi_0(t) \sim (1 - \lambda_\pi)^{-2}, \quad t \to \infty.$$

(b) *The distribution density for τ as $t \to 0$ has the form*

$$p_\tau(t) = \left[(1 - \lambda_\pi)\lambda_0(t) + \int_0^\infty \lambda_0(x)\lambda_0(x+t)dx + \lambda\right] \cdot (1 + o(1)), t \to 0. \tag{13}$$

Let $\pi_0(t)$ be monotone near zero, where $\pi(t) \sim At^{-p}$, $0 < p < 1$; let $\pi_0(t) < \varphi(t)$, where φ is a smooth function, $\int_0^\infty \varphi(t)dt < 1$, $\varphi(t) \sim ct^{-1-\theta}$, $t \to \infty$, $0 < \theta < 1$. Then

$$p_\tau(t) = O(\lambda_0(t)) \quad \text{as} \quad t \downarrow 0.$$

The time-magnitude self-exciting model. The self-exciting model is frequently considered on the time-magnitude space as follows (see, e.g., SAICHEV and SORNETTE, 2004): each event t_i (both when a main or a cluster one) is ascribed a random magnitude m_i. The m_i are independent for different t_i and have identical distributions with density $p(m)$. The generation of clusters is that described above, the only difference being that an event (s, m) generates a cluster with rate $q(m)\pi(t-s)$. It can be assumed without loss of generality that $\int q(m)p(m)dm = 1$. This normalization preserves statements 1, 3 for the self-exciting process (t, m) as well, independent of the choice of $p(m)$ and $q(m)$. The function $\lambda_0(t)$ as given by (12) then corresponds to the cluster rate when averaged over magnitude m. For purposes of seismology, $p(m)$ corresponds to the normalized Gutenberg- Richter law, $p(m) = \beta e^{-\beta(m-m_0)}$, $m > m_0$ while $q(m) = e^{\alpha(m-m_0)}(1-\alpha/\beta)$ is proportional to the size of the cluster that has been triggered by an event of magnitude m.

4. *General Point Processes and the Unified Scaling Laws*

The universal scaling law according to Corral. Let us consider waiting time τ between two consecutive seismic events of magnitude $m > m_c$ in area G. According to Corral's hypothesis (see CORRAL, 2003a,b), the distribution of the normalized variable τ: $\tau\lambda$ with $E\tau \cdot \lambda = 1$ is independent of the choice of the area and cutoff magnitude m_c. The probability density function of τ in this case has the form (1), i.e., $\lambda f(\lambda x)$, with a universal/unified function f. The term "unified" will be used if the hypothesis holds for any subarea of a fixed region G. Experiments which test (1) in CORRAL (2003b) concern both the Earth as a whole and smaller or larger areas of it. One can always select such areas in which seismicity is weakly interdependent, e.g., the regions of Spain and California which are tested by CORRAL (2003b) probably belong to this category. Our next statement is valid for general stochastic point processes for which τ has a probability density. We show that the class of the universal function f is very narrow if there are at least two seismogenic areas with independent seismicity (in stochastic sense).

Statement 4. *Assume that there are two regions G_1 and G_2 with independent stationary sequences of events $N_i(dt)$ and the normalized variable $\tau : \tau\lambda$ with $\lambda = 1/E\tau$, has the same probability density function $f(x)$ for $N_1(dt)$, $N_2(dt)$ and $N_1 + N_2$. If $f(x) < cx^{-\theta}$, $0 < \theta < 1$ for small x, then $f(x) = \exp(-x)$.*

Note that the process $N_1 + N_2$ corresponds to seismic events in the area $G_1 \cup G_2$. The exponential distribution of $\tau\lambda$ is realized in the case of the Poisson model of seismicity. This model is very crude because of the time-space clustering of events. Therefore Corral's universal scaling law cannot exist in the whole range of time. Note that outside of the aftershock's time zone $\tau \cdot \lambda > 0.05$, the empirical universal function f (see (2)) looks as $cx^{\gamma-1}\exp(-x/a)$ with $\gamma \neq 1$. Thus the clustering can not be the only obstacle for Corral's universal scaling law.

The unified law according to BAK *et al.* (2002). In contrast to Corral, the original version of the unified law (see BAK *et al.*, 2002) has to deal with τ variable, $\bar{\tau}_L$, for a *random* box area B_L of size L. The box population $\{B_L^{(i)}\}$ forms a square lattice covering of the seismogenic area G. If $\tau_i = \tau(B_L^{(i)})$ is the value of τ appropriate to the box $B_L^{(i)}$, then one has

$$\bar{\tau}_L = \{\tau_i \quad \text{with probability} \quad \lambda_i/\lambda,\}$$
$$\lambda = \sum \lambda_i$$

where $\lambda_i = \lambda(B_L^{(i)})$ is the rate of events of $m > m_c$ in $B_L^{(i)}$. BAK *et al.*'s scaling law means that the distribution of $\bar{\tau}_L\bar{\lambda}$ for a suitable constant $\bar{\lambda} = \bar{\lambda}(m_c, L)$ is independent of the choice of the lattice covering, L and m_c. The scaling factor $\bar{\lambda}$ can be found from the equation $\bar{\lambda}E\bar{\tau}_L = 1$. Hence

$$\bar{\lambda} = \lambda/N_L,$$

where N_L is the number of boxes $B_L^{(i)}$ with $\lambda_i \neq 0$. The value of $\bar{\lambda}$ is proportional to the original scaling factor $\tilde{\lambda} = 10^{-bm_c}L^d$ by BAK *et al.* (2002) where d is the box dimension of events of $m > m_c$ in G. This fact arises from the Gutenberg-Richter law and the definition of d resulting in $\log \lambda = c - bm_c$ and $\log N_L = -d \log L(1 + o(1))$ as $L \downarrow 0$. We will use the factor $\bar{\lambda}$ because the relation for N_L is inaccurate and the parameter d is debatable.

Usually, actual seismicity is non-homogeneous in space. If that is the case, we must consider the following.

Statement 5. *The unified laws by* BAK *et al. and by* CORRAL *in area G cannot exist simultaneously if the number of distinct values, k_L, among $\{\lambda(B_L^{(i)})\}$ depends on L, e.g. $k_{L_1} \neq k_{L_2}$ for two L only.*

However, a Poisson model with homogeneous rate in space and time is an example where the two laws exist simultaneously. The next statement shows that space homogeneity of seismicity is essential for the existence of unified BAK *et al.*'s law.

Statement 6. *Assume that $\gamma = \{B_L^{(i)}\}$ is the covering of $G = G_1 \cup G_2$, where G_1 and G_2 are non-intersecting subareas of G composed of elements of γ.*

If the random variables $\bar{\tau}_L\bar{\lambda}$, related to G_1, G_2 and G, have the same distribution, then the scaling factors $\bar{\lambda}$ for G_1 and G_2 are equal, $\bar{\lambda}(G_1) = \bar{\lambda}(G_2)$.

The simplest conclusion from this is the following

Corollary. *Suppose that the unified law by* BAK *et al. holds for area G and the rate of events $\lambda(\widetilde{G})$ continuously depends on subarea $\widetilde{G}$. Then λ is homogeneous, i.e. $\lambda(\widetilde{G}) = c|\widetilde{G}|$, where $|\widetilde{G}|$ is the area of $\widetilde{G}$.*

5. Conclusion

We have presented a theoretical analysis of the probability density function (PDF) of interevent time, τ, in a stationary point process. We show that if there are at least two seismogenic areas with independent seismicity, then a universal PDF. of $\tau\lambda$, so that $E\tau\lambda = 1$, must necessarily be exponential. The condition seems to be very natural in practice. Therefore Corral's unified scaling law (1) cannot exist in the entire range of time. In fact, the empirical data by CORRAL (2003b, 2004) supports the conclusion showing a universal behavior of PDF of $\tau\lambda$ only outside of the aftershock's time zone, i.e., for the time range $\tau\lambda > 0.05$. We show also that the original unified scaling law by BAK *et al.* (2002) can exist under nonrealistic conditions of homogeneity of seismicity rate (see Corollary for details).

Another aspect of this work is related to the asymptotic behavior of PDF of τ near 0 and ∞ in the general Poisson cluster model. The small-scale behavior of τ mimics the rate of cluster events, i. e., the Omori law for the case of seismicity. The asymptotics near ∞ are exponential due to the Poisson model of main seismic events. It is these asymptotics which essentially make the empirical PDF of $\tau\lambda$ universal when plotted on a log-log scale. In contrast to Corral, the unified empirical law by BAK *et al.* has a power asymptotics near ∞. According to CORRAL (2003a), it is a result of power-law tails in the distribution of space-time rate of seismicity.

Qualitatively the power and the exponential asymptotics are expected (see BAK *et al.*, 2002). A new element here is the explicit form of these asymptotics in connection with the general Poisson cluster model. This result can be used for analysis and parameterization of the universal part of PDF of $\tau\lambda$. According to CORRAL (2003b), the $\tau\lambda$ distribution is well represented by $f(x) = cx^{\gamma-1}\exp(-x/a)$ for $x = \lambda\tau \geq 0.05$. In the cluster model the parameter $1/a$ can be treated as the fraction of main events among all seismic events, λ^*/λ. The estimate $a = 1.23$ derived by CORRAL (2003b) yields $a^{-1} = 80\%$. At the same time the estimates of λ^*/λ are not so stable. The main events in Italy are 60% among the $m \geq 3.5$ events (MOLCHAN *et al.*, 1996). KAGAN (1991) found $\lambda^*/\lambda = 40\%$ for $m \geq 1.5$ using the CALNET catalog for the period 1971–1977. Variability of λ^*/λ versus stability of $1/a$ in CORRAL (2003a) is an interesting point for further analysis of universality of the tail probability of $\tau\lambda$.

In terms of the cluster model the factor $x^{\gamma-1}$ is missing in the formula for f. The factor may be replaced (see (7)) by a factor of the type $\exp(-cx^{1-\theta})$, if the aftershock rate decays as a power function $t^{-1-\theta}$, $0 < \theta \leq 1$; the factor degenerates to a constant for $\theta > 1$. Consequently, it remains an open question as to what is the physical meaning of γ.

Acknowledgements

I am very grateful to M. Shnirman for a helpful discussion and to D. Vere-Jones and I. Zaliapin for useful comments. This work was Partially supported by RFBR (grant 05-0564384 and by the EU 6-th Framework STReP project "E2-C2 extreme events: causes and consequences".

References

BAK, P., CHRISTENSEN, K., DANON, L., and SCANLON, T. (2002), *Unified Scaling Law for Earthquakes*, Phys. Rev. Lett. *88*, 178,501.

CORRAL, A. (2003a), *Local Distributions and Rate Fluctuations in a Unified Scaling Law for Earthquakes*, Phys. Rev. E *68*, 035102(R).

CORRAL, A. (2003 b), *Time-increasing Hazard and Increasing Time until the Next Earthquake*, ArXiv: cond-mat/0310407.

CORRAL, A. (2004), *Universal Local Versus Unified Global Scaling Laws in the Statistics of seismicity*, ArXiv: cond-mat/0402555.

DALEY, D. J. and VERE-JONES, D. *An Introduction to the Theory of the Point Processes*. Vol.1: *Elementary Theory and Methods*, 2-nd ed. (Springer, New-York, 2003).

FELLER, W. *An Introduction to Probability Theory and Its Applications*. Vol. II. (John Wiley and Sons, Inc., New York 1966).

HAWKES, A. G., and ADAMOPOULOS, L. (1973), *Cluster Models for Earthquakes—Regional Comparisons*, Bull. Int. Stat. Inst. *45*: 3.

KAGAN, Y. (1991), *Likelihood Analysis of Earthquake Catalogues*, Geophys. J. Int. *106*, 135–148.

MOLCHAN, G., KRONROD, T., DMITRIEVA, O., and NEKRASOVA, A. (1996), *Seismic Risk Oriented Multiscale Seismicity Model: Italy*, Comput. Seismol. *28* (in Russian).

OGATA, Y. (1986), *Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes*, Math. Seismol. *1*, 228–281, Inst. Statist. Math.

SAICHEV, A. and SORNETTE, D. (2004), *Anomalous Power-law Distribution of Total lifetimes of Aftershock Sequences*, ArXiv: physics/0404019.

UTSU, T., OGATA, Y., and MATSU'URA, R. S. (1994), *The Centenary of the Omori Formula for a Decay Law of Aftershock Activity*, J. Phys. Earth *43*.

VERE-JONES, D. (1970), *Stochastic models for earthquake occurrence*, J. Roy. Statist. Soc. B32.

Appendix

Proof of Statement 1

We are going to find the asymptotic of $P(\tau > t)$ as $t \to \infty$ using (6). Obviously that

$$B := \int_0^\infty P(N_0(t+du) > 0, N_0(u,u+t) = 0) \le \int_0^\infty P(N_0(t+du) > 0) = \int_t^\infty \lambda_0(u)du = o(1).$$

The limit for the expression under the exp sing in (6) is

$$C := \int_0^\infty P(N_0(u,t+u) > 0)du \to \int_0^\infty P(N_0(u,\infty) > 0)du =$$

$$= E\int_0^\infty \mathbf{1}_{N_0(u,\infty)>0}\,du = E\int_0^{\tau_{cl}} du = E\tau_{cl}.$$

Here, τ_{cl} is the cluster duration in N_0 and $\mathbf{1}_A$ is indicator function of an event A, i.e.

$\mathbf{1}_A = 1$, if A is true and $\mathbf{1}_A = 0$ otherwise.

It remains to substitute the resulting limits into the formula

$$P(\tau > t) = \exp\{-\lambda^*(t+C)\}[1+B]/(1+\Lambda). \tag{14}$$

We now are going to prove the second part of Statement 1. Let $E\tau_{cl} = \infty$. The asymptotic of C then calls for refinement. One has

$$P(N_0(u,t+u)>0) \le EN_0(u,t+u) = \int_u^{t+u} \lambda_0(v)dv.$$

If $\lambda_0(v) = c\theta(1-\theta)v^{-(1+\theta)}$ for $v >> 1$, then

$$C < \int_0^\infty du \int_u^{t+u} \lambda_0(v)dv = ct^{1-\theta}(1+o(1)), \quad t\to\infty,$$

as follows from L'Hospital's rule. Relation (7) stands proven.

Proof of Statement 2

The distribution density for τ can be found by differentiating (14):

$$f_\tau(t) = P\{N(0,t)=0\}\left[(1+B(t))^2\lambda^* - \dot{B}(t)\right](1+\Lambda)^{-1}.$$

One has $P(N(0,t)=0) \cong 1-\lambda t, \quad t \to 0$.

Due to condition (8),

$$B(t) = \int_0^\infty [1 - P(N_0(u,u+t)>0 \mid N_0\{u+t\}=1)]\,P(N_0(t+du)>0)$$

$$= \int_0^\infty \lambda_0(t+u)\,(1+o(1))\,du = \Lambda(1+o(1)), \quad t\to 0$$

and $\quad (1+B(t))^2\lambda^*(1+\Lambda)^{-1} = \lambda\,(1+o(1))$.

At last

$$[B(t)-B(t+\delta]/\delta = [\int_0^\delta P(N_0(t+du)>0, N_0(u,t+u)=0) +$$

$$\int_\delta^\infty P(N_0(t+du)>0, N_0(u,u+t)>0, N_0(u-\delta,u)>0)]/\delta$$

$$= I_1 + I_2.$$

Due to (8) the first summand is

$$I_1 = \frac{1}{\delta}\int_0^\delta P(N_0(t+du)>0)\,(1-P(N_0(u,u+t)>0 \mid N_0\{u+t\}>0) = \lambda_0(t)(1+o(1)), \delta\downarrow 0.$$

The second summand is $I_2 = \int_\delta^\infty \delta^{-1}P(N_0(t+du)>0, N_0(u-\delta,u)>0)\times$

$$[1-P(N_0(u,u+t)>0 \mid N_0\{u+t\}=1, N_0(u=1)].$$

By (9) $I_2 = \int_\delta^\infty \delta^{-1}P(N_0(t+du)>0, N_0(u-\delta,u)>0)\,(1+o(1))$.

Therefore $\dot{B}(t) = [\int_0^\infty \lambda_0(u)\lambda_u(t)du + \lambda_0(t)](1+o(1))$, where $\lambda_u(t)$ is the rate of $N_0(\cdot)$ at moment $t+u$ given $N_0(du) > 0$.

To prove (10) it remains to substitute the resulting asymptotic expressions in $f_\tau(t)$.

Proof of Statement 3

It follows from the description of the cascade generation of N_0 that its rate $\lambda_0(t)$ satisfies the integral equation

$$\lambda_0(t) = \int_0^t \pi_0(x)\lambda_0(t-x)dx + \pi_0(t), \tag{15}$$

where $\pi_0(t)$ is the rate of rank 1 events. Iteration of (15) then yields

$$\lambda_0(t) = \pi_0(t) + \pi_0 * \pi_0(t) + \pi_0 * \pi_0 * \pi_0(t) + \dots$$

If one passes to the Laplace transform, $\lambda \to \widehat{\lambda}$, then both relations for $\lambda_0(t)$ are reduced to the form

$$\widehat{\lambda}_0(s) = \widehat{\pi}(s)/(1-\widehat{\pi}(s)).$$

Let $\pi(t)$ be monotone near 0 and ∞. Assume also that $\pi(t)$ behaves like a power law: $\pi_0(t) \sim c_0 t^{-p}$, $t \ll 1$ or $\pi_0(t) \sim c_1 t^{-1-\theta}$, $t \gg 1$, where $0 < p, \theta < 1$. In that case the use of the Tauberian theorems (see FELLER, 1996), Ch. 13 and Ch. 17, 12) yields conclusions of the form $\lambda_0(t)/\pi(t) \to$ const as $t \to 0$ or $t \to \infty$, respectively.

We now prove (13). Consider the rate of a pair of events in an N_0 cluster: $\lambda_0(u,v) = P(N_0(du) = 1, N_0(dv) = 1)/(du\,dv), u < v$. Recalling that this is a cascade generation of N_0, the states u and v in N_0 can be derived in two ways. One is when u and v have no common parent except $t = 0$; the second is when u and v have a common parent z in the first generation (a state of rank 1). If the common parent z for u and v has rank $r > 1$, then the probability of that event will be of order $O((dz)^2 du\,dv)$, which is negligibly small compared with $O(dz\,du\,dv)$. This consideration leads to the following equation for $\lambda_0(u,v)$:

$$\lambda_0(u,v) = \lambda_0(u)\lambda_0(v) + \int_0^u \pi_0(z)\lambda_0(u-z, v-z)dz, \quad u < v. \tag{16}$$

Put $a_t(u) = \lambda_0(u, u+t)$, $b_t(u) = \lambda_0(u)\lambda_0(u+t)$, then (16) gives

$$a_t(u) = b_t(u) + \pi_0 * a_t(u),$$

whence

$$\begin{aligned} a_t(u) &= b_t(u) + b_t(u) * (\pi_0 + \pi_0 * \pi_0 + \pi_0 * \pi_0 * \pi_0 + \dots) \\ &= b_t(u) + b_t(u) * \lambda_0(u). \end{aligned} \tag{17}$$

We are interested in the conditional rate in a N_0 cluster:

$$\lambda_u(t) = \lambda_0(u, u+t)/\lambda_0(u).$$

Substituting a_t and b_t in (17) one has

$$\lambda_u(t) = \lambda_0(u+t) + \int_0^u \lambda_0(x)\lambda_0(x+t)\lambda_0(u-x)dx/\lambda_0(u).$$

It remains to substitute that expression in (10). One has

$$p_\tau(t) = [\lambda_0(t) + \int_0^\infty \lambda_0(u)\lambda_0(u+t)du \cdot (1+\Lambda) + \\ + \lambda(1+\Lambda)]/(1+\Lambda)\cdot(1+o(1)), \quad t \to 0.$$

However, $(1+\Lambda)^{-1} = 1 - \lambda_\pi$, so (13) is proved.

In order to have $p_\tau(t) = O(\lambda_0(t))$ as $t \to 0$, one has to show that $\lambda_0(u+t) < k\lambda_0(t)$ for small t. To do this, we demand $\pi_0(t) = c_1 t^{-\theta}, 0 < t < \varepsilon$ and $\pi_0(t) < \varphi(t)$. Here, φ is a smooth function, $\int_0^\infty \varphi(t) < 1$ and $\varphi(t) \sim ct^{-1-\theta}$, $t \gg 1$ with $0 < \theta < 1$; also, $\varphi = \pi_0$ for $t < \varepsilon$. Then

$$\lambda_\varphi = \varphi + \varphi * \varphi + \varphi * \varphi * \varphi \ldots$$

is a smooth function. One has $\lambda_0(t) < \lambda_\varphi(t)$ in virtue of (12), since $\pi_0 \leq \varphi$. One has $\lambda_\varphi(t)/\varphi(t) \to c$ as $t \to \infty$ from the power-law behavior of φ at ∞ (see Statement 3(a)). One also has $\lambda_0(t)/\pi_0(t) \to 1$ as $t \to 0$, hence $\lambda_0(t) \uparrow \infty$ as $t \to 0$. Consequently, $\max_{t>t_0} \lambda_0(t)$ will coincide with $\lambda_\varphi(t_0) = \lambda_0(t_0)(1+o(1))$ starting from some small t_0. Hence

$$\lambda_0(u+t) \leq \max_{v>t_0} \lambda_0(v) \leq \max_{v>t_0} \lambda_\varphi(v) \simeq \lambda_0(t_0)(1+o(1)).$$

Proof of Statement 4

Using (5) one has

$$p_\tau(t) = \frac{\partial^2}{\lambda \partial t^2} P\{N(0,t) = 0\}, \tag{18}$$

where λ is the rate of $N(dt)$ in the region. In virtue of (1),

$$p_\tau(t) = \lambda f(\lambda t).$$

Equation (18) and the initial conditions for $P\{N(0,t) = 0\} = u(t)$ having the form $u(0) = 1$ and $u'(0) = -\lambda$ specify $u(t)$ uniquely and yield $u(t) = \varphi(\lambda t)$, where

$$\varphi(t) = 1 - t + \int_0^t (t-s)f(s)ds. \tag{19}$$

Since $N_1(dt)$ and $N_2(dt)$ are independent,

$$P(N(0,t)=0)=P(N_1(0,t)=0)P(N_2(0,t)=0),$$

where $N=N_1+N_2$ is the sequence of events for $G_1\cup G_2$. It follows that for any $t>0$ one has

$$\varphi((\lambda_1+\lambda_2)t)=\varphi(\lambda_1 t)\varphi(\lambda_2 t).$$

or

$$\psi(t)=\psi(pt)+\psi(qt) \tag{20}$$

where $\psi(t)=\ln\varphi(t)$, $p=\lambda_1/(\lambda_1+\lambda_2)$ and $p+q=1$. Iteration of (20) yields for $p=q=1/2$

$$\psi(t)=\psi(\varepsilon_n t)/\varepsilon_n,\quad \varepsilon_n=2^{-n}$$

or

$$\frac{\psi(t)}{\psi(1)}=\frac{\psi(\varepsilon_n t)}{\psi(\varepsilon_n)}. \tag{21}$$

By $\varphi(0)=1$, $\varphi'(0)=-1$ we have $\psi(0)=0$ and $\psi'(0)=-1$. Using L'Hospital's rule we will have

$$\lim_{n\to\infty}\frac{\psi(\varepsilon_n t)}{\psi(\varepsilon_n)}=\lim_{n\to\infty}\frac{\psi'(\varepsilon_n t)t}{\psi'(\varepsilon_n)}=t.$$

By (21) one has $\psi(t)=-\alpha t$ or $\varphi(t)=\exp(-\alpha t)$. However, in that case $f(x)=\alpha^2 e^{-\alpha x}$ and $\int f(x)dx=1$, whence $\alpha=1$. Statement 4 is proven for $p=q=1/2$.

In the general case $p\neq 1/2$, the iteration of (16) yields

$$\psi(t)=\sum_{k=0}^{n} C_n^k\psi(\varepsilon_{k,n}t)$$

where $\varepsilon_{k,n}=p^k q^{n-k}$.

Similarly to the above, one has

$$\psi(\varepsilon_{k,n}t)=\psi(\varepsilon_{k,n})(t+\delta_{kn})$$

with $\delta_{kn}=o(1)$ as $n\to\infty$.

Using (19) and the a *priori* bound

$$f(x)<cx^{-\theta},\quad 0<\theta<1,\quad 0<x<\varepsilon$$

it is easy to show that

$$|\delta_{k,n}|<k_t\cdot[\max(p,q)]^{n(1-\theta)}.$$

Therefore we have $\psi(t) = \alpha t$ again, because the $\delta_{k,n}$ are small uniformly in k. The proof of Statement 4 is complete.

Proof of Statement 5

Let us prove Statement 5 by contradiction. We assume that the BAK *et al.* and the CORRAL laws hold for area G. Consider a box covering $\{B_L^{(i)}\}$ of G. Using the notation given in section 4, we have $p_i(t)$ for the distribution of $\tau_i = \tau(B_L^{(i)})$ and $\bar{p}(t)$ for that of $\bar{\tau}_L$. By the definition of $\bar{\tau}_L$ one has

$$\sum_{i=1}^{n} \lambda_i p_i(t) = \lambda \bar{p}(t), \quad \lambda_i > 0, \quad \lambda = \sum \lambda_i, \tag{22}$$

where $\lambda_i^{-1} = E\tau_i$. By Corral's hypothesis one has

$$p_i(t) = \lambda_i f(\lambda_i t) \tag{23}$$

and

$$\bar{p}(t) = \bar{\lambda} \bar{f}(\bar{\lambda} t), \quad \bar{\lambda} = \lambda / n. \tag{24}$$

The moments of order $0 < \alpha \leq 1$ for $p_i(t)$ are finite because $E\tau_i = \lambda_i^{-1} < \infty$. By (23) and (24) one has

$$E\tau_i^\alpha = \lambda_i^{-\alpha} m_\alpha, \quad E\bar{\tau}_L^\alpha = \bar{\lambda}^{-\alpha} \bar{m}_\alpha,$$

where $m_\alpha = \int x^\alpha f(x)dx$ and $\bar{m}_\alpha = \int x^\alpha \bar{f}(x)dx$.

Using (22) we come to the following relation:

$$\sum_1^n \lambda_i (\lambda_i n)^{-\alpha} = \lambda^{1-\alpha} \bar{m}_\alpha / m_\alpha,$$

The same relation holds for the other box covering of G with parameters $\{\tilde{\lambda}_j, j = 1, \ldots, m\}$; note that $\sum \tilde{\lambda}_j = \lambda$.

As a result one has

$$\sum_1^n \lambda_i (\lambda_i n)^{-\alpha} = \sum_1^m \tilde{\lambda}_j (\tilde{\lambda}_j m)^{-\alpha}, \quad 0 < \alpha < 1, \quad \lambda_i > 0, \quad \tilde{\lambda}_j > 0,$$

or

$$\sum_1^r \lambda_i k_i (\lambda_i n)^{-\alpha} = \sum_1^s \tilde{\lambda}_j \tilde{k}_j (\tilde{\lambda}_j m)^{-\alpha},$$

where $\lambda_1 < \ldots < \lambda_r$ and $\tilde{\lambda}_1 < \ldots < \tilde{\lambda}_s$, $\sum k_i = r$ and $\sum \tilde{k}_j = m$.

We have here the equality of two analytical functions of the argument α, when α member of $\in (0, 1)$. Therefore they are equal for all complex numbers α. But in that case $r = s$ and $\lambda_i n = \tilde{\lambda}_i m$, $i = 1, \ldots, r$. Therefore the number of distinct values among $\{\lambda_i\}$ cannot depend on L. This is a contradiction.

Proof of Statement 6

Let (22, 24) correspond to the covering $\{B_L^{(i)}\}$ of the area $G = G_1 \cup G_2$. The left-hand part of (22) contains terms related to the areas G_1 and G_2. Therefore (22, 24) lead to the following relation:

$$\lambda(G_1)\bar{p}_1(t) + \lambda(G_2)\bar{p}_2(t) = \lambda\bar{p}(t), \tag{25}$$

where $\bar{p}_i(t)$ corresponds to G_i and $\lambda = \lambda(G_1) + \lambda(G_2)$. By assumption relation (24) holds for $\bar{p}$ and for $\bar{p}_i$ with $\bar{\lambda} = \bar{\lambda}(G_i)$. Going over from (25) to the moments of order $0 < \alpha < 1$, we will come (see proof of Statements 5) to the equation:

$$\lambda(G_1)\bar{\lambda}_1^{-\alpha} + \lambda(G_2)\bar{\lambda}_2^{-\alpha} = \lambda\bar{\lambda}^{-\alpha}, \quad 0 < \alpha < 1, \tag{26}$$

where $\bar{\lambda}_i = \bar{\lambda}(G_i)$. By repeating the arguments with analytical extension of (26), we conclude that

$$\bar{\lambda}_1 = \bar{\lambda}_2 = \bar{\lambda}.$$

Let us prove the Corollary. We will consider two box coverings of G, i.e. $\{B_L^{(i)}\}$ and $\{B_{L_1}^{(j)}\}$, with $L_1 = L/2$, where L is small. One has $\bar{\lambda}(B_L^{(i)}) = \lambda(B_L^{(i)})/N_i$, where N_i is the number of boxes $B_{L_1}^{(j)}$ covering $B_L^{(i)}$ and heaving $\lambda(B_{L_1}^{(i)}) > 0$. By Statement 6, $\bar{\lambda}(B_L^{(i)}) = c(L)$. Therefore $\lambda(B_L^{(i)}) = c(L)N_i$ where $N_i \leq 4$, i.e., $\lambda(B_L^{(j)}) > 0$ admits only four non-zero values: $c, 2c, 3c$ and $4c$. Suppose $\lambda(B_L^{(i)}) \neq$ const., then we can find two adjacent boxes, say B_1 and B_2, so that $R = \lambda(B_1)/\lambda(B_2) < 1$, $R > 0$. Let us consider a continuous shift $\varphi(t)$ of $\{B_L^{(i)}\}$ such that $\varphi(1)B_1 = B_2$ and $\varphi(1)B_2 = B_3 \in \{B_L^{(i)}\}$. Then the ratio $R(t) = \lambda(\varphi(t)B_1)/\lambda(\varphi(t)B_2)$ is a continuous function with possible values i/j where $i, j = 0, 1, \ldots, 4$. Therefore $R(t) = R(0)$ for $0 \leq t \leq 1$, i.e. $R(1) = R(0)$. But $R(0) = i/j$, $i < j$ and $R(1) = j/m$, i.e. $j^2 = i \cdot m$. This equation has the unique solution: $j = 2$, $i = 1$ and $m = 4$. We can repeat our arguments for B_2 and B_3 again. However now we know that $i = 2$ and $j = 4$. Therefore one has $m = 8$, that is impossible because $m \leq 4$. Hence $\lambda(B_L^{(i)}) =$ const.

(Received April 21, 2004, accepted June 5, 2004)