Tectonophysics, 211 (1992) 85–98
Elsevier Science Publishers B.V., Amsterdam

On relations between earthquake population and asperity population on a fault

Alexander A. Gusev

Institute of Volcanology, Petropavlovsk-Kamchatsky, Russia

(Received March 18, 1991; revised version accepted October 18, 1991)

ABSTRACT

Gusev, A.A., 1992. On relations between earthquake population and asperity population on a fault. In: T. Mikumo, K. Aki, M. Ohnaka, L.J. Ruff and P.K.P. Spudich (Editors), Earthquake Source Physics and Earthquake Precursors. *Tectonophysics*, 211: 85–98.

Data were compiled on the trend of teleseismic short-period ($T \approx 1.4$ s) peak amplitude and spectral level vs M_0. For $\log M_0 > 26$, slopes (in log–log scale) of these trends were estimated as $b = 0.35$ and $\beta = 0.39$, respectively. For a similar trend of seismogram peak factor (peak to rms amplitude ratio) the slope estimate is $p = 0.13$.

$\beta = 0.39$ disagrees with the ω^{-2} spectral model but can be explained by the multiasperity fault model of Gusev (1989) if the typical asperity size $2R_a$ is assumed to grow slowly with M_0. If $R_a \propto M_0^{\delta}$, then $\beta = 1/3 + \delta$, and the empirical estimate of δ is 0.06. This value generally agrees with the f_{max} vs M_0 trend revealed recently. The last is supposed to reflect the R_a vs M_0 trend as well.

$p = 0.13$ is fully incompatible with a Gaussian process record model (predicting $p \approx 0.03$) and indicates a peak distribution of a heavy-tailed type. Assuming a power-law distribution for amplitudes of individual pulses (each one produced by failure of a single asperity) that add up to the observed record, we estimated the exponent α of the power law to be about 2.3. This may indicate that $\alpha \approx 2.3$ for the distribution of stress drop values of individual asperities. This α value agrees reasonably with $\alpha \approx 2$ found in Gusev (1989) from near-field data.

Theoretical α values are estimated for two hypothetical regular hierarchical asperity structures on a fault: a grid-like structure, giving $\alpha = 2$, and a clustered structure giving $\alpha \geqslant 1$. To obtain these estimates, we assumed a near-critical mode of overcoming barriers of successive scales during rupture propagation. Comparison with empirical data shows no contradiction, and suggests that some hierarchical asperity structure, probably a grid-like one, actually exists on natural faults.

Introduction

The short-period magnitude vs seismic moment relationship is one of the simplest and most reliable pieces of evidence that must be explained by any theory trying to describe radiation from an earthquake source. Generally, the same is true with respect to less reliable estimates of a similar trend of short-period spectral level vs seismic moment. Nevertheless no systematic explanation of these trends was proposed up to now. We shall try to suggest such an explanation based on the multiasperity fault model suggested recently by this author (Gusev, 1986, 1989). The main idea of this model, developed along the lines of Das and Kostrov (1983, 1986) and Boatwright (1988), is the successive failure of small asperities during earthquake rupture propagation. Comparing theory with near-field observations, typical parameters of asperities were estimated to be: size $2R_a = 0.5–1$ km, (local) stress drop $\Delta\tau = 300–800$ bar, filling factor (fraction of fault area covered by asperities) $k_f = 2–10\%$. The asperity stress drop distribution was close to the power-law (Pareto) type, with exponent α near 2. The multiasperity fault model provides a more or less coherent explanation of average trends and other features

Correspondence to: A.A. Gusev, Institute of Volcanic Geology and Geochemistry, Petropavlovsk-Kamchatsky 683006, USSR.

of near-field short-period data, including peak acceleration, acceleration spectral levels and spectral shapes. Hence, it is only natural to apply the multiasperity fault model to teleseismic data.

The first question we shall try to answer is whether teleseismic observations can be reasonably explained by this model. We shall show that such is indeed the case, and in particular that the assumption of a power-law type of distribution for asperity stress drop agrees with the data. This conclusion is of immediate relevance to the problem of dynamics of seismic rupture propagation.

Fukao and Furumoto (1985) have discussed in much detail the problem of stochastic propagation and termination of seismic rupture along a fault, including explanation of a power-law source size distribution, mainshock–aftershock magnitude difference and other evidence. Based on all this evidence they proposed the idea of a hierarchical, multiscaled structure of seismic ruptures and of a hierarchy of grid-like barriers on a fault which constitute boundaries that stop these ruptures. They applied an approach similar to the renormalization group approach. They found that in the 1D case, the probability density of the width of linear barriers can be described by a power function with exponent 2. In terms of distribution this means a power law with exponent 1. This result actually is valid for the 2D case as well. Gusev (1989) replaced the mentioned linear barriers of variable width by chains of asperities of variable strength, and showed that this strength should be power-law distributed, with the exponent α near to 2. In order to understand the possible meaning of the observed α values we shall develop these results, and arrive at some rather preliminary conclusions on the possible structure of the asperity distribution over a fault.

On amplitude and spectral trends of earthquake source radiation in the 0.5–1 Hz frequency band

In order to compare theoretical and observed properties of short-period earthquake source radiation one needs to summarize relevant observational results. To do so one can compile the published data. Such a compilation was carried out recently by the author (Gusev, 1991); we shall shortly cite the relevant points here, and also add some new information. Revised intermagnitude relationships were constructed using the M_0 scale as the fundamental one. In particular, trends of short-period magnitude $m_{\mathrm{PV}}^{\mathrm{SKM}}$ of the Soviet ESSN service and its "western counterpart" m_b^* (Koyama and Zheng, 1985) or equivalently $\hat{m}_b$ (Houston and Kanamori, 1986) vs M_0 were found; they were parallel, differing by a constant in the average:

$$m_b^* = m_{\mathrm{PV}}^{\mathrm{SKM}} - 0.18 \tag{1}$$

At $26 < \log\, M_0 < 30$, the trend is log-linear:

$$m_{\mathrm{PV}}^{\mathrm{SKM}} = b\, \log\, M_0 + C \tag{2}$$

and values $b = 0.35$ and $C = -2.75$ were determined rather reliably. Although no formal error value was given, the real accuracy of b is about 0.02. One can compare this b value to the earlier estimates of Koyama and Shimada (1985), equivalent to $b = 0.40$, and of Houston and Kanamori (1986), equivalent to $b = 0.37$. Both the linearity of $m_{\mathrm{PV}}^{\mathrm{SKM}}$ vs log M_0 and the particular value of b present important information on the source spectral structure. To prevent confusion we must mention that standard m_b of NEIC is incomparable with $m_{\mathrm{PV}}^{\mathrm{SKM}}$ or m^*db because, according to accepted routine, for large enough M_0 it is actually not based on the true maximum of a record. Let us recall now that at a given distance,

$$m_{\mathrm{PV}}^{\mathrm{SKM}} = \log(A_{\mathrm{sp}}/T_{\mathrm{sp}}) + C_1 \approx \log\, A_{\mathrm{sp}} + C_2 \tag{3}$$

where A_{sp} is the amplitude of the short-period P wave record, that is of source P wave radiation filtered by a certain narrow band filter with a central frequency of about 0.7 Hz, and where T_{sp} is the visual period of the record which can be considered M_0-independent (Houston and Kanamori, 1986; Boore, 1986). In other words, the response of a short-period instrument combined with an absorption filter produces a signal which can safely be considered narrow-band. Therefore, for A_{sp} one can write:

$$A_{\mathrm{sp}} \propto M_0^{0.35} \tag{4}$$

Additional data on this point are average trends of frequency band PV magnitudes determined by Zhbrykunov and Zhbrykunova (1974) following ideas of K.K. Zapolskii. These magnitudes were determined for many earthquakes recorded by octave-bandwidth filters added to a conventional seismograph ("ChISS station"). Each of two cutoffs of these filters was formed by a fourth-order analog filter. For each of three bands with central frequencies of 0.67, 1 and 1.5 Hz, the equivalent b value for a log M_0 range of 26.5–28.5 is equal to the common figure of 0.35. The original data were presented in the m_{PV} scale; we reduced them to the M_0 scale using the m_{PV} vs M_0 relation from (Gusev, 1991), which is near to the standard ones.

In Gusev (1991) data were also compiled on the trend of short-period (around 0.7 Hz) spectral level S_{sp} versus M_0, and it has been shown that for β in:

$$S_{sp} \propto M_0^{\beta} \tag{5}$$

teleseismic P wave data give a β value of ~ 0.39. This value practically coincides with the spectral trend of California earthquakes (at 0.7 Hz) deduced by Papageorgiou and Aki (1985) from local data. Earlier teleseismic estimates of β are: 0.37 according to trends in Gusev (1983), 0.50 after Koyama and Zheng (1985), 0.45 after Houston and Kanamori (1986) and 0.41–0.52 for subduction zone earthquakes after Zhuo and Kanamori (1987). Data of Hartzell and Heaton (1985) indicate $\beta = 0.33$ to 0.40 at $M = 7.25$ to 8.75, but show saturation for larger events; we believe that this saturation was produced by some peculiarity of their data set. To estimate b, common catalogue data on m and M_0 (hundreds of events) were used for a log M_0 range of 25–30; thus $b = 0.35$ can be thought as the reasonable world average; however, only a limited data set (tens of events) could be compiled for the estimation of β, and its accuracy is lower. We believe that the real accuracy can be represented by the estimate range of $\beta = 0.375$ to 0.42 (related primarily to the variations between data sets, and only secondarily to the statistical errors proper). Therefore we consider the difference between the empirical β and the theoretical $\beta = 1/3$ for the ω^{-2} model as being real; the same can be concluded from the β estimates cited earlier.

Based on the values $b = 0.35$ and $\beta = 0.39$, one can estimate the trend of the peak factor PF of the teleseismic record. Use of this parameter, defined as the ratio of peak to rms-amplitude, has an important advantage—it enables one to decouple effects related to signal statistics from spectral trends as such, which is desirable. Assume for sufficiently large earthquakes:

$$PF \propto M_0^{p} \tag{6}$$

Then one can estimate p in the following way. From Parseval's theorem one may write the rms-amplitude as:

$$A_{rms} = S_{sp}(\Delta f/d)^{0.5} \tag{7}$$

where d and Δf are effective duration and bandwidth, respectively. One can assume Δf as constant for a short-period record. As for d, for large enough M_0 (log $M_0 > 26$) the assumption of similarity gives $d \propto M_0^{1/3}$; combining this with (5) and (7) gives:

$$A_{rms} \propto M_0^{\beta - 1/6} \tag{8}$$

At the same time, $A_{peak} \propto M_0^{b}$ (4), hence:

$$PF = A_{peak}/A_{rms} \propto M_0^{b - \beta + 1/6} \tag{9}$$

and:

$$p = b + 1/6 - \beta \tag{10}$$

For the cited b and β, $p = 0.13$. The pair (β, p) presents independent information on the spectral trend and record statistics.

Now we shall shortly discuss the meaning of the value $\beta = 0.39$ (a more detailed discussion is outside the scope of the present paper). Note that the ω^{-2} spectral models (Aki, 1967; Brune, 1970) predict $\beta = 1/3$ for large enough magnitudes. The descriptive $\omega^{-\gamma}$ model of Thatcher and Hanks (1973) predicts $\beta = 0.39$ at $\lambda = 1.83$. This last model, however, lacks clear physical interpretation. The multiasperity fault model predicts (Gusev, 1989, eq. 19) the following relation for the intermediate spectral level between common corner frequency f_0 and the characteristic frequency f_a (usually, $f_a = 3$ to 6 Hz) related to asperity failure time $T_a \approx 1/f_a$:

$$S_{sp} \propto S^{1/2} R_a \tag{11}$$

where S is the source area and $2R_a$ is asperity size, assumed similar for all asperities in a given source. If R_a is independent of magnitude, similarity gives $\beta = 1/3$ again. However, recent publications (Umeda, 1981; Umeda et al., 1984; Aki, 1986, 1987; Faccioli, 1986) indicate that the upper cutoff frequency $f_{\max}$ of the acceleration spectrum decreases with magnitude. This trend of $f_{\max}$ can be considered as the direct indication of the growth of R_a with magnitude. (The idea that $f_{\max}$ is related to the source spectrum is not common, however: many people relate the $f_{\max}$ phenomenon in general to effects of the frequency-dependent absorption in the medium.) Assuming the form of the R_a vs M_0 trend to be:

$$R_a \propto M_0^{\delta} \tag{12}$$

then $\beta = 1/3 + \delta$. As for the number of N asperities in a particular source, it is proportional to the source area S and inversely proportional to area S_1 associated with each asperity. Let $S_a = \pi R_a^2$, then:

$$S_1 = S_a/k_{\mathrm{f}} = \pi R_a^2/k_{\mathrm{f}} \tag{13}$$

where k_{f} is the filling factor, or fraction of fault area covered by asperities. Assuming k_{f} to be constant and $S \propto M_0^{2/3}$, we obtain $S_1 \propto R_a^2$ and:

$$N \propto SR_a^{-2} \propto M_0^{2/3-2\delta} \tag{14}$$

If $b = 0.39$, $\delta = 0.06$. This estimate of δ is preliminary, but it does not contradict the observed $f_{\max}$ trends. The fact that the multiasperity fault model provides a common explanation to two independent pieces of data looks hopeful.

Interpretation of the trend of the peak factor of a short-period record

To understand the meaning of the observed p value we need some theoretical reference point. The simplest is the Gaussian-process model of a seismic record, in which an apparently random seismic record is represented as a realization of a (maybe modulated) Gaussian process (Gusev, 1979, 1983; Hanks and McGuire, 1981). It is well known that for the simplest version of this model —modulation by boxcar of duration d—the peak behaves as:

$$PF = (\ln f_{\mathrm{r}} d)^{0.5} \tag{15}$$

where f_{r} is some representative (average) frequency. This relation is non-linear on a log–log scale, but can be approximated as linear. Hanks and McGuire (1981) give $p = 0.03$ as a useful approximation for accelerograms. Since the only relevant value is the number of peaks, this value remains valid for a short-period teleseismic record as well, because the numbers of peaks are comparable. In both cases, only sufficiently large magnitude events were considered so that the $d \propto M_0^{1/3}$ relation was thought to be valid. This means, roughly, $M > 5$ for accelerograms at a distance of 25 km, and $M > 6.5$ for teleseismic records.

The value $p = 0.03$ is strongly different from the observed $p = 0.13$, and this is a good reason to reject the Gaussian model for the teleseismic case. Lack of non-linearity of the magnitude trend is an additional argument. The Gaussian model is doubtful also for accelerograms (Gusev, 1989). In a search for a better theory one has to find a model which conserves the random nature of a record but ascribes larger probability to high peaks ("heavy-tailed" distribution). These properties are secured by some special kind of pulse process model, which represents a record as a sum of similarly shaped pulses whose onset times are random (forming a Poisson-type point process) and whose amplitudes are also random, and obey some reasonable heavy-tailed distribution law. In such a model, the power (and Fourier) spectral shape is determined by pulse shape only, whereas peak statistics depend on pulse amplitude statistics and also on pulse density.

Generally, the choice of a particular family of distribution laws to describe pulse statistics is somewhat arbitrary, but the condition of a linear p vs M_0 trend (on a log–log scale) identifies distributions with a power-law tail as the most convenient class. Between these, the power-law or Pareto distribution:

$$\Pr(x > x') \equiv 1 - F(x) = P(x) = \begin{cases} x^{-\alpha}: & x > 1 \\ 1: & x \leqslant 1 \end{cases} \tag{16}$$

is the simplest and will be assumed to govern pulse amplitudes. Strictly speaking, $P(x)$ is a complementary cumulative distribution function; we shall refer to it simply as a distribution function for brevity.

This model was successfully applied recently to accelerograms which were supposedly produced by rupture of a multiasperity fault (Gusev, 1989). In that case, individual acceleration pulses were short enough and pulse overlapping was negligible. In such a simple case, pulse amplitude distribution merely coincides with record peak distribution. With short-period teleseismic records, the situation is more complicated. Breaking of a single asperity produces a displacement step of duration T_a. This step passes through a medium absorption filter that cuts off nearly all energy above 0.8–0.9 Hz, and also through an instrument filter. Their joint effect produces essentially a narrow-band filter, with the central frequency close to 0.7 Hz. It will convert displacement steps of duration T_a into pulses of typical duration equal to an inverse filter/record bandwidth of about 2 s. Because of the narrow-band nature of a record, it may be considered also as a filtered (inverted) accelerogram; this enables one to relate record peaks to asperity stress drops, as was done in Gusev (1989) with respect to near-field accelerograms. Now note that individual pulses, as they are modified by the described filter, will strongly overlap. Their number (or, rather, rate) will not, however, be very large (as needed to safely use asymptotic results of the theory of probability). Hence, to study the proposed model in detail we need numerical simulations. Some preliminary results can be obtained, however, employing general properties of distributions with power-law tails (Feller, 1971, Chapters 8 and 14).

The first important result is that the asymptotic behaviour of a sum of independent variates with a common power-law distribution with exponent α depends substantially on the actual α value. When $0 < \alpha < 2$, this sum converges to some asymptotic law having a power-law tail with the same α. However at $\alpha \geqslant 2$ the asymptotic law is Gaussian (as it must be if variance is finite). The convergence rate in the second case depends on α, and it is very low for α values near to 2. As for the maximum among variates distributed with a power-law tail with exponent α, its median will be distributed also with a power-law tail of the same exponent α for any α. What does all this mean for our problem?

Let α be smaller than 2, and assume for a while that pulses fully overlap, i.e. their relative time shift is zero. Then distribution of amplitude of their sum will have a power-law tail with the same α. Otherwise, if pulses are so rare that there is no overlapping, peak amplitude will be distributed as a maximum among power-law variates, that is, again with a power-law tail of the same α. We can try to interpolate between these extreme cases and to guess that in case of $\alpha < 2$, the distribution of peaks will always have a power-law tail with the same α. If, on the other hand, $\alpha \geqslant 2$, in the case of a zero time shift the amplitude of the sum will be Gaussian-distributed *asymptotically*, that is in the limit of a very large number of pulses, and the difference between the actual and the Gaussian distribution will increase as the number of pulses becomes smaller and as the α value approaches 2. The same will be generally true for randomly phased pulses: the peak distribution law will be determined by the α value and actual multiplicity m of overlapping pulses.

To estimate the order of magnitude of m, let $M = 8$, $S = 8000$ km^2, $S_a = 2$ km, $k_f = 5\%$, $d = 50$ s; then $N = 200$, pulse rate is 4 s^{-1} and, for a pulse duration of 2 s, $m = 8$. Simple numerical estimates show that for α between 2 and 3, the tail of the distribution of a sum of 10–15 variates is not near to that of a Gaussian distribution, but can be approximately described instead by a power law with some effective exponent α_e (which is always greater than the actual α for pulses).

Therefore we can propose the following preliminary interpretation of the observed value α_p of the exponent of the distribution law for record peaks. If α_p is below 2, then the "true" α (for pulses) is equal to α_p, and if α_p is above 2 then the "true" α is bracketed between 2 and α_p. To apply this scheme to real data we must connect α_p with p which is known from observation. To do so we need a theoretical formula for p for a pulse process. This can easily be done. Let A_1 be the mean or median peak of single pulse, and d_1 its total duration. Pulse rms-amplitude (based on duration d_1) can be written for this case as:

$$A_{\mathrm{rms},1} = cA_1 \tag{17}$$

with a particular constant c. The pulse energy is, by definition:

$$E_1 = d_1 A^2_{\text{rms},1} \tag{18}$$

For N randomly phased pulses, the energy is NE_1, and the rms-amplitude of a segment of process with duration d and containing N pulses is:

$$A_{\text{rms}} = (NE_1/d)^{0.5} = A_1(cd_1)^{0.5}(N/d)^{0.5} \tag{19}$$

For non-overlapping pulses with a power-law amplitude distribution, the median value for the maximum among N amplitudes is (see Gusev, 1989):

$$A_{\text{p}} \approx A_1 N^{1/\alpha} \tag{20}$$

In the case of $0 < \alpha < 2$, this formula remains true for overlapping pulses because any overlapping has only minor effect on the strongest pulse (see Feller, 1971). For the peak factor we obtain:

$$PF = A_{\text{p}} = A_{\text{rms}} = (cd_1)^{-1/2} d^{1/2} N^{1/\alpha - 1/2} \tag{21}$$

In the case of $\alpha \geqslant 2$ and strong overlapping, we assumed the record peak distribution to have a power-law tail with exponent α_{p}; this immediately gives (analogue of (20)):

$$PF \propto n^{1/\alpha_{\text{p}}} \tag{22}$$

where n is the number of independent extrema. Substituting $N(M_0)$ from (14) into (21) and assuming $n \propto d \propto M_0^{1/3}$, for these two cases we obtain:

$$p = \delta + (2/3 - 2\delta)/\alpha_{\text{p}} - 1/6 \tag{23}$$

or:

$$p = 1/3\alpha_{\text{p}} \tag{24}$$

respectively. Substituting $p = 0.13$ and $\delta = 0.06$ gives $\alpha_{\text{p}} = 2.29$ from the first formula. Thus α is above 2 and the second formula may be more adequate. It gives $\alpha_{\text{p}} = 2.56$, therefore α ranges between 2 and 2.6. We shall use below $\alpha = 2.3$ as a preliminary result of interpretation of the peak factor.

We believe that the α value recovered from a short-period teleseismic displacement record corresponds to the amplitudes of acceleration pulses produced by failure of individual asperities. As was shown by Gusev (1989), these amplitudes are proportional to local stress drops $\Delta\tau$ of these asperities (averaged over the asperity area). Therefore our results indicate that the distribution of $\Delta\tau$ of individual asperities can have a power-law tail with $\alpha \approx 2.3$. This is the main result of this section.

Note that $\Delta\tau = \tau_{\text{c}} - \sigma_{\text{f}}/k_{\text{f}}$, where τ_{c} is actual asperity strength and σ_{c} is (residual) fault friction. If friction were zero, the α value would describe the asperity strength distribution. During the theoretical treatise below we shall set $\sigma_{\text{f}} = 0$ for simplicity.

Theoretical estimates of α

As we have found empirical estimates of exponent α, we could try to make some conclusions on an earthquake fault structure if we know the α values for different models of this fault. In this section we shall try to derive such theoretical α estimates. Generally speaking, the accurate theoretical way of determining the α value for a realistic multiasperity fault model could consist of the following steps: (1) assume some reasonable rules governing the relative asperity location over the fault and their strength distribution, including the distribution law (e.g. power law) for a single asperity, two-point correlation function, three-point correlation and so on; (2) for a numerical model, simulate a fault according to the rules; (3) load the fault by remote stress and simulate and record a history of its failure; (4) repeat (1)–(3) for different sets of rules; and (5) sort out a class of histories which look realistic, resulting in a judgement of the actual fault structure.

Realization of such a program is far outside the scope of the present paper; rather we shall carry out a very preliminary study based on simple examples. Instead of specifying random patterns by their correlation properties we shall assume certain deterministic structures. Since we will employ essentially the ideas of the renormalization group approach, our structures will be hierarchical. In particular, we shall discuss the hierarchical pattern of grids and the hierarchical pattern of clusters. In both cases, initial construction is a multiasperity fault regularly covered by asperities of similar size and shape; their strength however varies.

Hierarchy of grids

Fukao and Furumoto (1985) proposed that barriers on a fault are linear features which form a hierarchy of grids of different strengths: the larger is the cell of a grid, the stronger are the constituent linear barriers. Resistance of linear barriers to load was specified as "barrier height", equivalent to barrier width. The problem of barrier strength was not discussed at all, and we may believe that in the described model, local barrier strength (critical stress) is the same for all barriers, narrow and wide. Alternatively, one may assume varying barrier resistance to be produced by different barrier critical stress at constant width, or, in a general case, by some combination of both.

One of the main ideas of Fukao and Furumoto (1985) is that earthquake rupture grows by self-similar stages, and that instantaneous rupture size at each stage coincides with one or more cells of the corresponding grid. The perimeter of each cell is assumed to be broken in a near-critical mode; this explains why rupture neither always stops at some small cell size nor always propagates to infinity, but instead stops with a nearly constant probability at each successive stage producing a realistic frequency–magnitude relationship. Note that the "maximum magnitude" frequency–magnitude relationship which is specific for the near-source area of a characteristic earthquake (Wesnousky et al., 1984), suggests that in many cases rupture growth may be near to supercritical up to its final size. If this mode is typical, the results presented below can describe certain marginal cases.

Now let us try to obtain some information on the α value (eq. 16) from crack mechanics considerations. We shall proceed as follows: (1) we shall assume that any fault patch fails near-critically, and shall determine the proportion of fault surface which must be occupied by barriers for patches of a given size; (2) using this result for patches of different sizes, we shall determine the statistical distribution of fault surface elements with respect to their strength; and (3) we shall apply the resulting exponent of the power law to the strength values of asperities. To start with, we consider the condition of failure for a square cell of some definite side $2L$ surrounded by a square strip barrier of width $D \ll L$, whose strength is much above the strength of its surroundings. The state of the perimeter barrier is critical when relevant conditions of failure are fulfilled, e.g. it may be loaded by a stress equal to the barrier strength σ_c. We need to relate such "external" parameters as D, L and remote stress σ_∞ with certain "inner" parameters which determine the conditions of failure. Our main problem is that we do not know either the adequate failure criterion, or, hence, the relevant inner parameter. In crack mechanics, various failure criteria are in use, and with each of them different results can be obtained. Actually, the situation greatly simplifies, as will soon be seen.

Dropping all factors of about unity, for a Barenblat-Dugdale type model of a square shear crack of size L (friction is set to zero for simplicity) we may write (see e.g. Rice, 1980):

$$\sigma_c^2 a \approx \sigma_\infty^2 L \approx K^2/\mu \approx G \approx \sigma_c h \qquad (25)$$

where a is the cohesion zone width, K is the stress intensity factor for equivalent Griffiths crack (at $a \to 0$), μ is the shear module, G is energy consumption rate, and h is the critical displacement defined by the condition that shear traction at fault wall drops from σ_c to zero (or to friction stress) when shear displacement overgrows h.

Now we can consider different cases, when different inner parameters are assumed constant. In particular, (1) if a is a material constant (independent of L), $\sigma_c \propto L^{1/2}$; (2) if σ_c is a material property (elastoplastic model), $a \propto L$; (3) if h is a material constant, then $a \propto L - 1$ and $\sigma_c \propto L$; (4) if K (or G) is a material constant, then all cracks of a size below some L_0 stop and all cracks of a size above L_0 run to infinity, thus this assumption is inadequate for the problem under study.

To relate these general considerations to our problem we shall make the somewhat arbitrary assumption that in the critical state, the a value coincides with barrier width D. Of course D cannot be smaller than a; but the opposite is possible; in other words, the strong strip-barrier can fail not as a whole but in steps (one substrip after another). We shall assume that our barriers are marginally strong and fail in one step. Thus $a = D$, and for a given L:

$$\sigma_c^2 D \approx \sigma_\infty^2 L \approx \sigma_c h \approx G = \text{constant} \tag{26}$$

Consider the generalized case $D \propto \sigma_c^y$ covering all three reasonable versions discussed ($y = -1$ gives $h = \text{const}$, $y = 0$ gives $D = \text{const}$, and $y = \infty$ gives $\sigma_c = \text{const}$). Now (26) gives:

$$\sigma_c \propto L^{1/(2+y)} \propto S_c^{1/2(2+y)} \tag{27}$$

where S_c is cell area, and also:

$$D \propto L^{y/(2+y)} \tag{28}$$

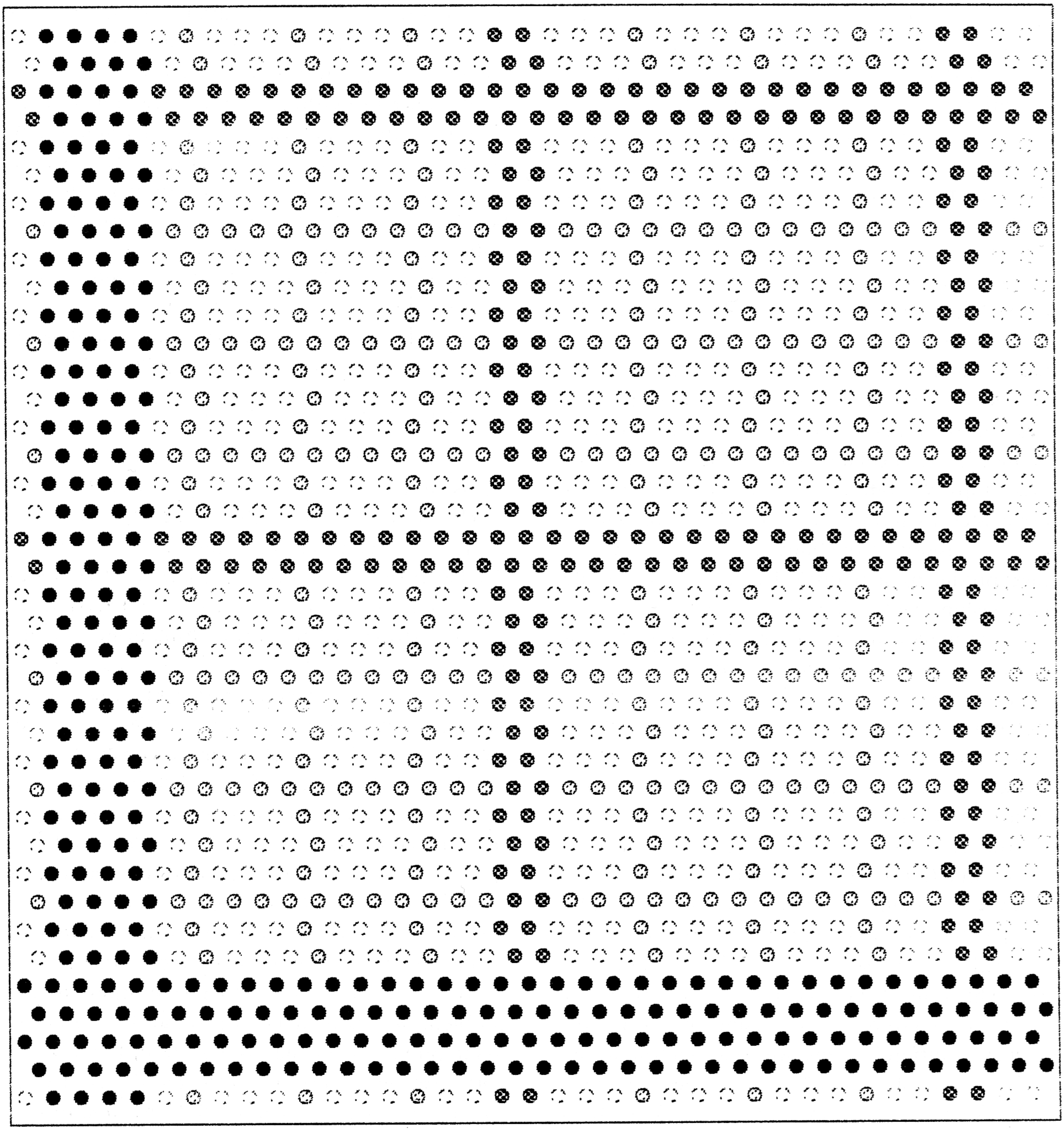

Fig. 1. Hierarchical grid-like multiasperity structure. Case of $D \propto L^{0.5}$, $y = 2$ and $\tau \propto L^{0.25}$. Shading of an asperity symbol reflects asperity strength.

For the value of area $S_b = 4DL$ of the perimeter barrier of the cell, this gives:

$$S_b \propto L^{(2+2y)/(2+y)} \propto S_c^{(1+y)/(2+y)} \tag{29}$$

Let us try now to determine the fraction of fault area of which the strength is near to σ_c. Following again the reasoning of Fukao and Furumoto (1985) and of Gusev (1989), we assume that patches of size S_c form a grid covering the whole surface. Then taking in account that each patch "owns" half of its perimeter, for the ratio of an area of barriers of strength σ_c to the area of patches we obtain:

$$p(\sigma_c) = S_b/2S_c \propto S_c^{-1/(2+y)} \propto \sigma_c^{-2} \tag{30}$$

for any $y < \infty$.

As was explained in detail in Fukao and Furumoto (1985) and in Gusev (1989), there is wide evidence indicating that there exists a whole hierarchy of barrier grids on natural faults, and one can believe that in grids with larger cell size their perimeter barriers are of greater strength. Hierarchically, this makes discrete σ_c values to form a geometric series, whereas fractions of fault area which correspond to these σ_c values follow eq. 30. Note that the value of p is an equivalent of probability. We must emphasise that eq. 30 covers cases of constant h, constant a, and all intermediate ones but the case of constant σ_c. This case is specific; it means that all barriers are of the same local strength, leading to a delta-like asperity strength distribution. We consider this case as contradicting observations and shall not mention it further. We can conclude that eq. 30 must be valid for a rather wide class of fault models.

Now we have to pass from the probability value $p(\sigma_c)$ to the (complementary) distribution law $P(\sigma_c)$. Note that if the probability density of x is a power law with exponent ϵ, the corresponding complementary distribution law $p(x) = Pr\ (x' > x)$ will be a power-law with exponent $\epsilon - 1$ because of integration. Meanwhile, if the probability of values of x is represented by atoms distributed along the x-axis according to geometric series, and the weights of these atoms follow a power law with exponent ϵ, a (step-like) distribution law will have the same exponent ϵ. The proof is trivial, and was given in Gusev (1989). Our case is evidently the second one; thus $P(\sigma_c) \propto \sigma_c^{-2}$. Therefore, the assumption of hierarchical linear barriers failing in critical mode can be realized by a fault strength distribution of a special kind—a power law with exponent 2.

In order to apply this result to a multiasperity fault we may follow the reasoning of Gusev (1989) and assume that the average local fault strength of σ_c can be approximately realized by a weak surface covered by asperities of strength τ, with filling factor k_f, so that $\sigma_c = k_f \tau$. A strong strip or linear barrier is replaced now by one or more rows of asperities (see Fig. 1 for example). The results concerning the σ_c distribution can be transferred to the τ distribution giving at last:

$$Pr(\tau' > \tau) = P(\tau) \propto \tau^{-2} \tag{31}$$

which is the result we sought for. It coincides with a similar result of Gusev (1989) which was derived, however, for a less general case: all barriers were assumed to contain only one row of asperities.

Hierarchy of clusters

To create an example of clustered hierarchical structures we shall follow the general approach, due to Fournier d'Albe, described in Mandelbrot (1982) in connection with fractal stellar clusters. The construction procedure for fractal-type structures which is applicable in this case starts with fixing the lower fractal limit (smallest scale or grain size) and repeating construction (or elimination) of substructures, multiplying their size, step by step, up to infinitely large ones. Our particular construction mode is as follows. At step 0, minimal strength τ_0 is ascribed to all asperities. A repeating pattern is chosen, which is an $M \times N$ box containing K black and $MN - K$ white cells. We cover all the surface by replicas of this pattern and ascribe strength $\tau_1 > \tau_0$ to all asperities marked black (forgetting their old strength value). Step 1 is ready. Then we enlarge the pattern up to the size $M^2 \times N^2$, so that old black cells are replaced by black blocks of $M \times N$ size. Covering the surface by replicas of the enlarged pattern we ascribe strength $\tau_2 > \tau_1$ to all asperities with "old" strength τ_1 covered by each black block. Step 2 is ready, and the way of induction for following steps seems clear. Figure 2 presents an illustration of the resulting construction for $M \times N = 2 \times 3$ and $K = 3$.

Therefore all asperities do not form a fractal pattern, but for their subset containing all asperities with a strength above some particular strength level, the pattern is fractal for some range of scales, and the higher this level, the wider is the range. An important parameter of our construc-

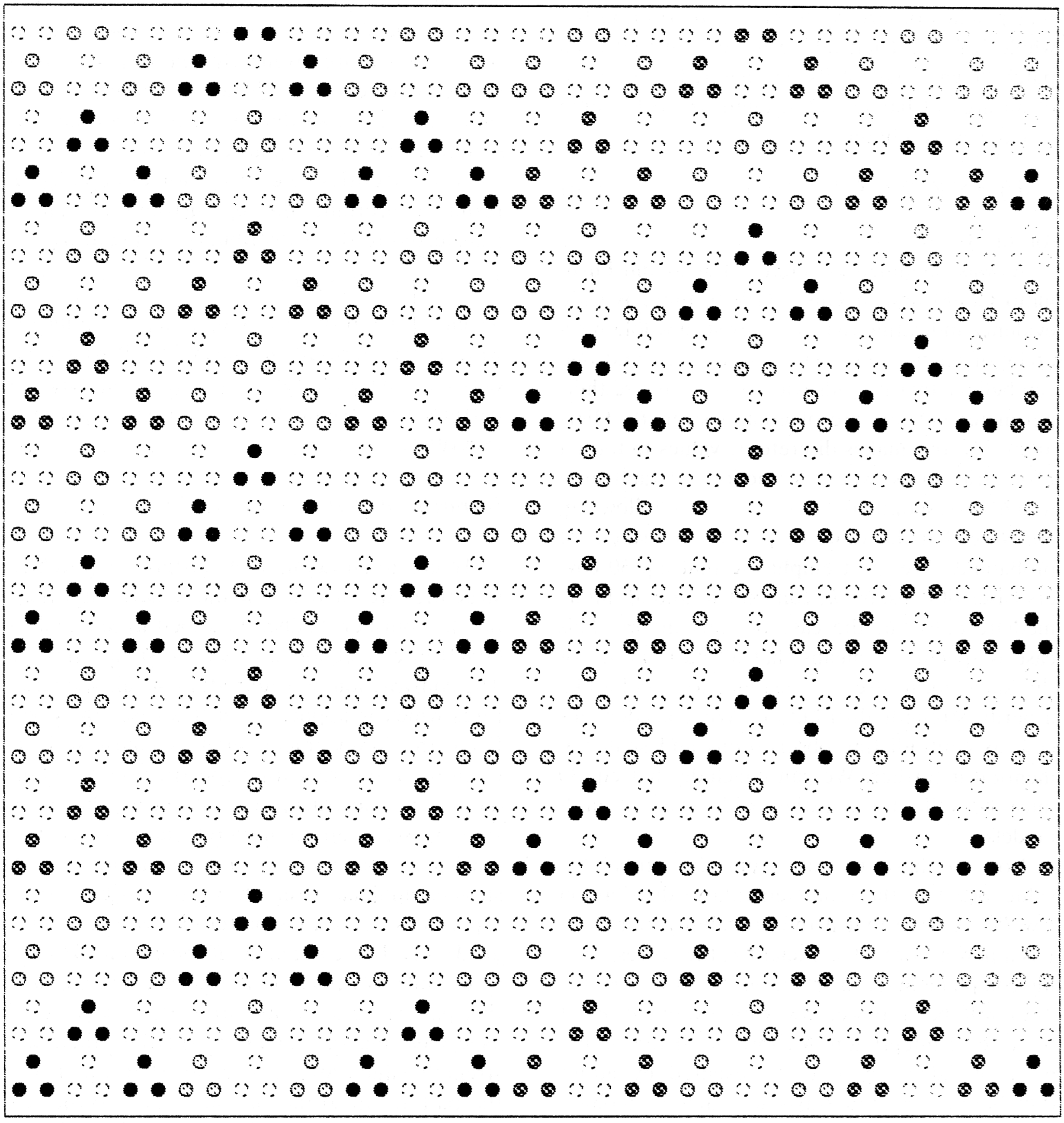

Fig. 2. Hierarchical clustered multiasperity structure. Example with $M \times N = 6$, $K = 3$ and $\kappa = 0.5$. Shading of an asperity symbol reflects asperity strength.

tion is $\kappa = K/MN$, which describes "rarefaction" of strong asperities at each step. In our example, $\kappa = 3/(2 \times 3) = 0.5$. The relative (in respect to step 0) density d_k of asperities considered at step k is therefore equal to κ^k. After examination of Figure 2 we may assume that for large enough loads, fault resistance will be localized in certain clusters, and gaps between them will easily yield. We can assume that during rupture growth over such a fault, these clusters will form temporary or final boundaries for this rupture. Consider some more or less isometric gap and note that it always contains inside itself some cluster of lower scale. Assume that this gap is occupied by rupture, and that the inner cluster fails in a near-critical mode.

One can roughly estimate the critical load over the cluster using the results of Das and Kostrov (1986) for a circular asperity of radius R_a in the centre of a circular cut of radius R. Average stress over this asperity (assumed to be near to the critical one σ_c) is:

$$\sigma_c = \sigma_\infty (R/R_a) \tag{32}$$

In a fractal structure, an equivalent of the R/R_a ratio is constant, so that $\sigma_c \propto \sigma_\infty$. (This result is in fact general and does not depend on our use of the particular model of Das and Kostrov, 1986.) The average number of asperities which actually resist in the central cluster is not easy to estimate; it is bracketed between the total number N_t of asperities in the cluster and the number of only the strongest asperities $N_{st} = N_t \kappa^k$. The first limiting case is improbable; it leads actually to constant strength of all asperities. The second limiting case seems to be probable, but with no guarantee until actual modelling. Considering this case as the limiting one we may note that it assumes the minimal number of resisting asperities and thus the maximum strength of them; therefore, this limiting case gives a lower bound $\underline{\alpha}$ for the value of α that is specific for the pattern under study. Let us determine $\underline{\alpha}$.

Since the average σ_c over the central cluster is constant, the total force loading the cluster is $N_t \sigma_c S_1$. It is distributed over the N_{st} strongest asperities and loads each of them with the average stress:

$$\tau = (N_t/N_{st}) \sigma_c k_f^{-1} = \sigma_c k_f^{-1} \kappa^{-k} \tag{33}$$

Since the average relative density, or "probability" of these asperities is κ^k, this gives $p(\tau) \propto \tau^{-1}$, leading, for already discussed reasons, to:

$$P(\tau) \propto \tau^{-1} \tag{34}$$

Therefore, $\underline{\alpha} = 1$ for a clustered structure of this particular kind. This lower bound estimate of α is preliminary, and must be confirmed by simulation. As for α proper, we found no way to get even a preliminary estimate; direct simulation is needed. One may conjecture, however, that this value will be < 2, because a clustered structure "uses" asperity strength for resistance less effectively than a grid structure.

The presented results describe particular models, and it is not clear whether they can be generalized. They may reflect, however, some general properties of hierarchical asperity distributions. One can suspect that adding randomness to the grid model never improves its resistance; thus, for more general grid models, $\alpha \leqslant 2$. Therefore, as a preliminary result we may guess that for clustered and grid-like hierarchical models in general, when a near-critical growth mode is assumed, α is bracketed between 1 and 2, with lower values for clusters and higher ones for grids.

Another problem is the relation of the discussed model to nature. Important evidence that supports an idea of grid-like structures is presented in Fukao and Furumoto (1985), but it is not conclusive and one can continue to consider all the construction as artificial. We can mention in this relation that purely random fractal models of topography demonstrated by Mandelbrot (1982) include many "lakes" (that is, depressions circled by some ridge) and the sizes of these lakes are power-law distributed. Mandelbrot even guesses that lakes cover the surface so that its remainder is of zero area (and of a fractal dimension between 1 and 2). This can mean that a certain grid of ridges of a strength function can be formed under rather weak assumptions.

Note that we consider here a fractal strength function. This point of view is not to be confused with another one, in which a fractal stress or stress drop function is assumed (von Seggern, 1981; Andrews, 1981).

General discussion of α estimates

Let us compare different estimates of the exponent α of the asperity stress drop distribution. Empirical data for the near field (Gusev, 1989) indicate that α is close to 2 based on the following observations ordered by their relative weight:

(1) $\alpha = 2$ provides good agreement between two empirical estimates of average/median $\Delta\tau$, one of which is based on the maximum value of an average Fourier acceleration spectrum and the other on the acceleration peak value.

(2) $\alpha = 2$ provides realistic (fast) growth of near-source peak acceleration with a magnitude in the range $M = 4$ to 6.

(3) $\alpha = 2$ provides independence of peak acceleration of fault distance up to distances near to the source width.

One can note, however, that all this evidence does not constrain α strongly, and in terms of a range, indicate, roughly, $\alpha = 1.7$–2.5.

Far field data presented in the first part of the present paper use fully independent information on α. They lead to the interval estimate $\alpha = 2.2$–2.6, which may be interpreted as $\alpha \approx 2.3$. As all these estimates are of limited accuracy, we may conclude that the empirical α value is near to 2–2.3.

Theoretical results were obtained for two particular deterministic asperity patterns, and one may doubt whether any generalization may be done. In both cases, some type of hierarchical structure was assumed, and the condition of a near-critical growth mode was applied to derive estimates of α. This value was 1 or greater for the studied clustered barrier structure, and 2 for the grid-like barrier structure. One can suspect that for an unstructured (delta-correlated) asperity distribution either α is equal to 1 or less, or the near-critical mode of growth normally cannot be achieved at all. As real asperity patterns can correspond either to a near-critical or a supercritical mode of growth, any theoretical estimate of the kind described actually provides the lower bound.

Now we can proceed to a comparison. The first observation is that ranges of empirical and theoretical estimates generally overlap. This is an implicit test of the whole approach (which is far from being a strict one), and this test is successful. The numerical correspondence of $\alpha = 2$ for a particular grid model and of empirical $\alpha = 2$ to 2.3 is of less importance. The studied grid model is somewhat artificial, and a lower value of α, say, $\alpha = 1.5$, can be expected in the model if randomness and discreteness will be accounted for, whereas $\alpha = 2$ to 2.3 may be the actual value in nature. The difference between these values can be produced by a mainly supercritical mode of growth of real ruptures. Nevertheless, our results generally agree with the idea of Fukao and Furumoto (1985) on the existence of a hierarchy of barrier grids on a fault surface.

Caputo and Console (1980) noted that distribution of empirical stress drop in two samples of earthquakes (for California and Japan) can be described by a power law, with exponents of 1 and 0.5, respectively. Inspection of corresponding figures shows that the tail part of the distributions is described only poorly, and greater values of α, say, $\alpha = 1.5$ to 2, can be needed to describe this tail. Though this information is not definite enough, it induces a general and interesting question: what kind of relation can one expect between the asperity stress drop distribution over a fault and the distribution of average (over rupture area) stress drop values of individual earthquakes. To derive such a relation we may reason as follows.

Assume the whole fault surface to be divided into patches of a size about 1 km, and consider the statistical distribution of all patches in respect to their average critical stress drop $\Delta\tau$. Choosing some threshold $\Delta\tau_0$ we may sort all patches into "asperities" with stress drop $\Delta\tau > \Delta\tau_0$, and "weak patches" with $\Delta\tau < \Delta\tau_0$. We may guess that nature does not keep the discussed distribution bimodal, because no reasons are seen for any natural threshold value. Much more likely is that the $\Delta\tau$ distribution is common for all patches, so that a threshold value is fully arbitrary. If so, we may assume the $\Delta\tau$ distribution of patches to be the power-law one with some exponent α. Since the average stress drop value is, in our representation, an average over several patches, its distribution will correspond to that of a sum of independent power-law variates, so that all results concerning such sums (mentioned above) are applicable. In particular, if α for asperity stress drops is less than 2, α for earthquake stress drops will be the same.

Conclusion

Application of the multiasperity fault model of Gusev (1989) to the interpretation of short-period teleseismic amplitude and spectral trends was successful. This lends additional support to this model and also enables one to compare an empirical asperity strength distribution with theoretical ones. The most important results of this study are:

(1) Slopes of log–log amplitude and spectral level vs M_0 for teleseismic short-period records (frequency band around 0.7 Hz) are near to $b = 0.35$ and $\beta = 0.39$, correspondingly, for sufficiently large earthquakes ($\log M_0 > 26$).

(2) The slope of the record peak factor (peak to rms-amplitude ratio) is near to $p = 0.13$, in pronounced contradiction with the Gaussian process record model (which predicts $p \approx 0.03$).

(3) The value $\beta = 0.39$ of spectral slope is in contradiction with $\beta = 1/3$ of the ω^{-2} model; it is shown to agree with predictions of the multiasperity model if one assumes slow growth of asperity size $2R_a$ with magnitude. If $2R_a \propto M_0^{\delta}$, then $\beta = 1/3 + \delta$ and empirical δ estimate is 0.06. This result generally agrees with the observed f_{max} vs M_0 trend.

(4) Interpretation of a peak factor trend can be based on the representation of a record as a random pulse process with a heavy-tailed pulse amplitude distribution. A simple assumption of a power-law distribution leads to a preliminary estimate of exponent α of this distribution; α is between 2 and 2.6, and $\alpha = 2.3$ is the preliminary point estimate.

(5) The same α value describes the distribution of average local stress drop values of individual fault asperities. This agrees reasonably with the near-field estimate of $\alpha = 2$ from Gusev (1989).

(6) The α value for strength and/or stress drop is determined for two deterministic hierarchical models of a multiasperity fault. The first model is the "discretized" analogue of the hierarchical grid model of Fukao and Furumoto (1985), which gives $\alpha = 2$. The second model is a hierarchical cluster model analogous to the stellar cluster model presented by Mandelbrot (1982), which gives $\alpha \geqslant 1$.

(7) General correspondence may be noted for α ranges from empirical data ($\alpha = 2$ to 2.3) and from theory ($\alpha = 1$ to 2). The difference may reflect the fact that in the theoretical derivation we assumed a near-critical mode of source growth corresponding to a Gutenberg-Richer type power-law moment–frequency relation, whereas actually the characteristic earthquake model and maximum magnitude (Wesnousky et al., 1984) moment–frequency relation may be more adequate, indicating a supercritical mode of growth.

References

Aki, K., 1967. Scaling law of seismic spectrum. J. Geophys. Res., 72: 1217–1231.

Aki, K., 1986. Origin of f_{max}. A preprint.

Aki, K., 1987. Magnitude–frequency relation for small earthquakes: a clue to the origin of f_{max} of large earthquakes. J. Geophys. Res., 92: 1349–1355.

Andrews, D.J., 1981. A stochastic fault model, 2. Time-dependent case. J. Geophys. Res., 86: 3867–3877.

Boatwright, J., 1988. The seismic radiation from composite models of faulting. Bull. Seismol. Soc. Am., 78: 489–508.

Boore, D.M., 1986. Short-period P and S wave radiation from large earthquakes: implications for spectral scaling relations. Bull. Seismol. Soc. Am., 76: 43–64.

Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes. J. Geophys. Res., 75: 4997–5009.

Caputo, M. and Console, R., 1980. Statistical distribution of stress drops and faults of seismic regions. Tectonophysics, 67: T13–T20.

Das, S. and Kostrov, B.V., 1983. Breaking of a single asperity: rupture process and seismic radiation. J. Geophys. Res., 88: 4277–4288.

Das, S. and Kostrov, B.V., 1986. Fracture of a single asperity on a finite fault: a model for weak earthquakes? In: S. Das, J. Boatwright and C.H. Scholz (Editors), Earthquake Source Mechanics. Am. Geophys. Union, Washington, pp. 91–96.

Faccioli, E., 1986. A study of strong motions from Italy and Yugoslavia in terms of gross source prosperities. In: S. Das, J. Boatwright and C.H. Scholz (Editors), Earthquake Source Mechanics. Am. Geophys. Union, Washington, pp. 297–310.

Feller, W., 1971. An Introduction to Probability Theory and its Applications (2nd ed.). Wiley, New York, N.Y., 669 pp.

Fukao, Y. and Furumoto, M., 1985. Hierarchy in earthquake size distribution. Phys. Earth Planet. Inter., 37: 35–148.

Gusev, A.A., 1979. Descriptive statistical model of short-period earthquake source radiation. Dokl. Akad. Nauk SSSR, 244: 544–548 (in Russian).

Gusev, A.A., 1983. Descriptive statistical model of earthquake source radiation and its application to short-period strong motion. Geophys. J.R. Astron. Soc., 74: 787–808.

Gusev, A.A., 1986. On the nature of short-period radiation of an earthquake source. In: Seismicity and Seismic Prediction in the Far East, Abstr. Conf., Petropavlovsk-Kamchatsky, p. 79 (in Russian).

Gusev, A.A., 1989. Multiasperity fault model and the nature of short-period subsources. Pure Appl. Geophys., 130: 635–660.

Gusev, A.A., 1991. Intermagnitude relationships and asperity statistics. Pure Appl. Geophys., 136: 515–527.

Hanks, T.C. and McGuire, R.K., 1981. The character of high-frequency ground motion. Bull. Seismol. Soc. Am., 71: 2071–2095.

Hanks, T.C., 1982. f_{max}. Bull. Seismol. Soc. Am., 72: 1867–1879.

Hartzell, S.H. and Heaton, T.H., 1985. Teleseismic time functions for large shallow earthquakes. Bull. Seismol. Soc. Am., 75: 965–1004.

Houston, H. and Kanamori, H., 1986. Source spectra of great earthquakes: teleseismic constraints on rupture process and strong motion. Bull. Seismol. Soc. Am., 76: 19–42.

Hwang, L.J. and Kanamori, H., 1989. Teleseismic and strong-motion spectra from two earthquakes in eastern Taiwan. Bull. Seismol. Soc. Am., 79: 935–944.

Koyama, J. and Shimada, N., 1985. Physical basis of earthquake magnitudes: an extreme value of seismic amplitudes from incoherent fracture of random fault patches. Phys. Earth Planet. Inter., 40: 301–308.

Koyama, J. and Zheng, S.H., 1985. Excitation of short period body waves by great earthquakes. Phys. Earth Planet. Inter., 37: 108–123.

Mandelbrot, B.B., 1982. The Fractal Geometry of Nature. Freeman, New York, 457 pp.

Papageorgiou, A.S. and Aki, K., 1983. A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion, 1. Description of the model. Bull. Seismol. Soc. Am., 73: 953–978.

Papageorgiou, A.S. and Aki, K., 1985. Scaling law of far-field spectra based on observed parameters of the specific barrier model. Pure Appl. Geophys., 123: 353–374.

Rice, J., 1980. The mechanics of earthquake rupture. In: A.M. Dziewonsky and E. Boschi (Editors), Physics of the Earth's Interiors. Proc. Int. School of Physics 'Enrico Fermi', Course 78, Italian Phys. Soc., pp. 555–649.

Thatcher, W. and Hanks, T., 1973. Source parameters of southern California earthquakes. J. Geophys. Res., 78: 8547–8576.

Umeda, Y., 1981. An earthquake source model with a ripple generating core. J. Phys. Earth, 29: 341–370.

Umeda, Y., Iio, Y., Kiroiso, A., Ito, K. and Murakami, H., 1984. Scaling of observed seismic spectra. Zisin, 37: 559–567.

Von Seggern, D., 1980. A random stress model for seismicity statistics and earthquake prediction. Geophys. Res. Lett., 7: 637–640.

Wesnousky, S.G., Scholz, C.H., Shimazaki, K. and Matsuda, T., 1984. Integration of geological and seismological data for the analysis of seismic hazard: a case study for Japan. Bull. Seismol. Soc. Am., 74: 687–708.

Zhbrykunova, N.A. and Zhbrykunov, V.Ya., 1974. On a maximum magnitude in P wave CnISS spectrum. In: Magnitude and Energy Classification of Earthquakes, 1. Nauka, Moscow, pp. 198–206 (in Russian).

Zhuo, Y. and Kanamori, H., 1987. Regional variation of the short-period (1–10 seconds) source spectrum. Bull. Seismol. Soc. Am., 77: 514–529.