ISSN 1028-334X, Doklady Earth Sciences, 2007, Vol. 417, No. 8, pp. 1216–1220. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © Yu.L. Rebetskii, 2007, published in Doklady Akademii Nauk, 2007, Vol. 416, No. 5, pp. 680–685.

GEOPHYSICS

Stressed State Corresponding to the Formation of Large-Scale Brittle Failure of Rocks

Yu. L. Rebetskii

Presented by Academician A.O. Gliko October 2, 2006

Received October 5, 2006

DOI: 10.1134/S1028334X07080181

Issue of relations between brittle failure and localized plastic flow, which are realized in the fault zones of the Earth's crust, is one of the most important problems in geomechanics. New data on values of natural stresses in seismoactive regions allowed us to distinguish the regularities of their distribution, which correspond to the possibility of the appearance of large-scale brittle failure. It was found that areas of medium and low levels of confining pressure are the most hazardous regions in terms of a strong earthquake. Here, a small part of the released energy is consumed for negotiating friction forces. Therefore, brittle failure is more efficient. In the regions of high rock pressure, the dissipation of elastic energy is realized as creep along the fault or cataclastic flow due to a large number of weak earthquakes.

INTRODUCTION

The majority of the models of earthquake sources are based on Reid's concept of the stage of earthquake preparation: an earthquake is preceded by the appearance of an obstacle (joint) to the displacement along part of the Earth's crust, which governs gradual increase in the actual stresses up to the limiting values. Within this concept, the preparation time and magnitude of an earthquake are directly related to the size of the region with anomalously high levels of deviator stresses.

This concept, however, did not yield a technology for the discrimination of regions of strong earthquake preparation. The successes in prediction are related to the magnitudes of earthquakes with high recurrence (M = 4–6). The unique case of prediction of an earthquake with M = 7.3 near Haicheng (China) in 1975 was followed by the catastrophic Tanshan earthquake (1976), which was not manifested in any forerunners.

NATURAL STRESSES IN ROCK MASSIFS

Failures in the field of earthquake prediction are related to the following fact. Many concepts about earthquake preparation were adopted from mechanics of construction materials. They do not take into account specific features of the structure of seismogenic regions (fractures) in the Earth's crust. Only recently were new data on the values of natural stresses obtained as a result of the development of the method of cataclastic analysis of seismological data on mechanisms of earthquake sources [1, 2]. This method allows us to estimate the stresses and parameters of strength of rock materials in their natural state.

The investigations of the parameters of natural stresses showed that the level of deviator stresses at a scale of n–10n km does not exceed 100n bars, while the effective intact cohesion between the rock massifs τ_f is 10n bars [3]. It was found that strong earthquakes occur usually in regions with low and medium levels of deviator stresses (Fig. 1a). The data in Fig. 1a also reflect the mutual correspondence of the maximal shear stresses τ and effective isotropic pressure p^* ($p^* = p - p_{fl}$, where p is pressure in solid rocks, and $p_{fl} > 0$ is fluid pressure in fissured-porous space) in the crustal regions where the stressed state is close to the limiting one (according to Coulomb).

INHOMOGENEITY OF STRESSES IN THE EARTHQUAKE PREPARATION REGION

The results of field studies and the theoretical analysis presented above were obtained a few years ago. They definitely contradicted the generally accepted concepts of the process of earthquake occurrence and required explanation. This explanation appeared after the analysis of the results of reconstruction of the stressed state preceding the catastrophic Sumatra–

Institute of Physics of the Earth, Russian Academy of Sciences, ul. Bol'shaya Gruzinskaya 10, Moscow, 123810 Russia; e-mail: reb@ifz.ru

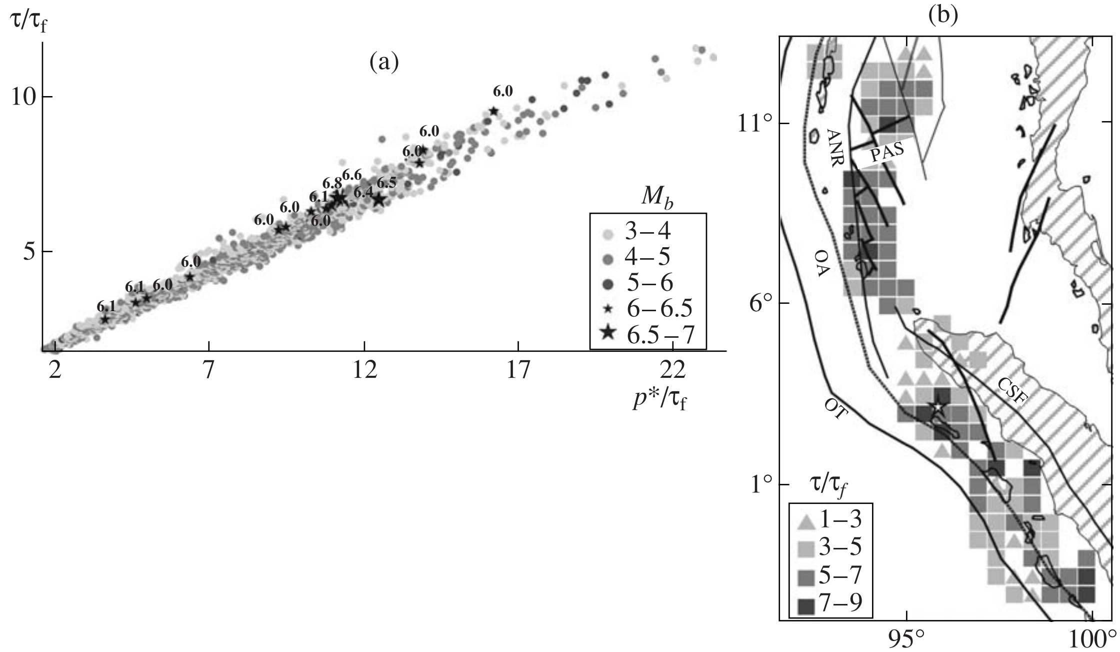

Fig. 1. Relation between the stress level and intensity of seismic radiation. (a) Sources of strong earthquakes in the parametric region of relative deviator and isotropic stresses in the Earth's crust of the Japanese Islands. (b) Field of relative values of maximal shear stresses on the western flank of the Sunda arc before the SAE. The asterisk denotes the onset of opening of the SAE source. (CSF) Central Sumatra Fault, (ANR) Andaman–Nicobar Fault; (OT) oceanic trough; (OA) oceanic arc; (PAS) pull-apart structures in the back-arc basin.

Andaman earthquake (SAE) on December 26, 2004. The source of this earthquake was spread over more than 1250 km to the north of the onset of its opening (Fig. 1b) and exceeded many times the linear scale of stress averaging based on reconstruction (the calculated parameters for seismological data from the Harvard University catalogue are averaged over the volume with a linear size of 50–100 km). Thus, we could see for the first time a stress field in the earthquake preparation region.

The SAE demonstrated that the character of the stressed state and, what is most important, the distribution of stress are significantly nonuniform along the future earthquake source (Fig. 1b). This type of nonuniformity in the Sunda arc region is caused by the active regions of the Andaman–Nicobar and Central Sumatra dextral strike-slip faults. The onset of opening of the earthquake source was located in the region of the maximal shear stresses (approximately 250 bar) near its boundary with the region of their lowest level (approximately 100 bar). In these regions, the level of the effective isotropic pressure, which forms an obstacle to the development of brittle failure, was equal to 420 and 150 bar, respectively. The length of the regions of increased and decreased levels of stresses along the subduction zone was equal to 200–250 and 300–400 km, respectively.

The SAE source was opened at the northern boundary of the region with a high stress level. The dynamic spreading of the source with release of seismic energy took place along the north-northwest direction, i.e., toward the region of lower stresses via the region of the high stress gradient. The seismological data showed that at the initial time moment equal to 50 s, which is presumably the time of propagation across the region of a high level of stresses, the velocity of fault spreading was low (2 km/s) at a low intensity of seismic radiation. The velocity sharply increased to 3 km/s later [5]. The fault region (320–350 km long) near the northern end of Sumatra is characterized by the greatest displacements (up to 20 m) and a magnitude of $M_w \approx 9$. Results of the reconstruction of stresses show that the high velocity of the front in this fault region was caused by the low level of friction forces that provided high efficiency of the earthquake.

RELATION BETWEEN BRITTLE FAILURE AND PLASTIC FLOW

Thus, the results obtained revealed the following fact: (i) strong earthquakes occur in the regions of medium and low levels of deviator and isotropic stresses (Fig. 1a); (ii) opening of the SAE source was directed to the northern region of a medium level of stresses and was not developed in the southern region of a high level of these stresses (Fig. 1b). In order to explain this fact, we had to address once more the experimental results pertaining to the destruction of rock samples. They suggest that both brittle failure and localized plastic flow can be developed in samples at a

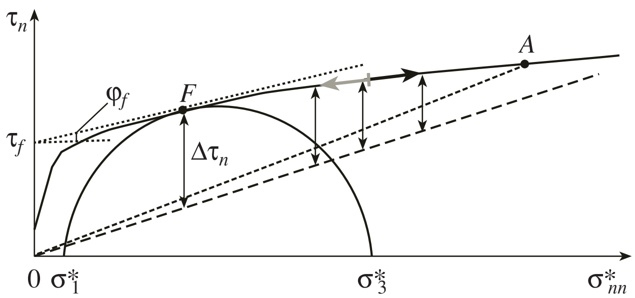

Fig. 2. Relation between the regions of brittle failure and plastic flow in the Mohr diagram. The solid line shows the intact strength of rocks; dashed (short and long) lines show minimal resistance of dry static friction for the existing fissures and kinematic friction realized during the motion of the walls. The dotted line corresponds to the determination of strength parameters ($k_f = \tan\varphi_f$ and τ_f) at point F. Vertical bars correspond to the values of stress drop at the corresponding points. The gray arrow determines the direction of variation in the stressed state at which brittle failure is more efficient. The black arrow shows the direction along which the value of seismic radiation decreases. Point A divides the diagram into the regions of brittle failure (left) and plastic flow (right).

certain level of isotropic pressure. Both mechanisms allow the rocks to release part of the energy accumulated in elastic deformations. The first type of brittle failure occurs at relatively low isotropic pressure and corresponding deviator stresses. At a high level of isotropic pressure, brittle failure gives way to a localized plastic flow.

Such an interrelation is caused by the following fact: the energy of elastic deformations released during brittle failure is consumed for the formation of plastic and microbrittle deformations near the end of the newly forming fault and the negotiation of friction forces during the sliding of its walls. The rest of the energy is spent on the growth of surface energy and fault and the generation of elastic seismic waves. The greater the level of the effective isotropic pressure p^*, the greater the part of the energy spent for the negotiation of friction forces. When the intensity of stresses increases, the efficiency of an earthquake, i.e., the ratio of the energy of seismic waves E_q to the total released energy, decreases:

$$\eta_q = \frac{E_q}{|\Delta E_e|} = \frac{0.5\Delta\tau_n}{\tau_n^0 - 0.5\Delta\tau_n}, \tag{1}$$

where τ_n^0 is the mean shear stress along the fracture, which acted before the formation of the fracture, and $\Delta\tau_n$ is the released mean shear stress.

Decrease in the efficiency of an earthquake is related not only to the increase in the energy needed to overcome the friction forces and finally transferred to the thermal energy, but also to the curvilinear form of the Mohr envelope (Fig. 2) that governs the decrease in the coefficient of internal friction k_f at a higher level of stress. In the samples, the destruction region is formed by the creep mechanism in the limiting case; i.e., we see a plastic flow, which is accompanied by microbrittle deformations. The ambiguity in the relations between brittle failure and stress level was missed by the authors of all known models of destruction.

Figure 2 shows the Mohr diagram corresponding to the experimental results of the destruction of rocks [6]. With consideration of the decrease in the effective strength parameters at large scales of averaging, the diagram also shows a qualitative pattern of the behavior of rock massifs. We can distinguish two characteristic lines. The curvilinear envelope characterizes the achievement of the limiting stressed state when new fractures and faults are formed. The rectilinear curve beginning from the origin of coordinates governs the minimal resistance of dry static friction in the preexistings. Brittle failure of the same rocks can occur in a wide range of maximal shear stresses τ, i.e., when the limiting state rather than the limiting stress τ is gained (this fact is frequently ignored). Brittle failure is possible in initially fractured rocks when the large Mohr circle crosses the dry (static) friction line (Fig. 2). Thus, the points characterizing the stress vector (τ_n, σ_n) on the plane of a developing fracture are located within the region bounded by this line and the limiting Mohr envelope.

When averaged with the linear size, the value of stress drop $\Delta\tau_n$ corresponding to the length of the fracture can be estimated graphically on the Mohr diagram using the concept about constant effective normal stresses σ_{nn}^* on the fracture plane before and after its formation (activation) [7]. The $\Delta\tau_n$ value related to the case of the limiting stressed state of the newly formed fracture corresponds to the length of the vertical segment, which connects the point of contact of the large Mohr circle of the upper limiting envelope with the line of resistance with kinematic friction k_k (if the displacement direction of fracture walls coincides with the direction τ_n^0):

$$\Delta\tau_n = \tau_f + (k_f - k_k)\sigma_{nn}^*. \tag{2}$$

Here, τ_f and k_f are the intact cohesion of rocks and the coefficient of internal friction, which are effective parameters corresponding to their scale of averaging for large volumes of rocks, and $\sigma_{nn}^* = \sigma_{nn} - p_{\text{fl}}$ is the mean effective stress normal to the fracture.

The analysis of the dependence of $\Delta\tau_n$ on the intensity of the stressed state in the brittle failure regions (Fig. 2) shows that the maximal $\Delta\tau_n$ values are observed in the region with medium values of deviator and isotropic stresses. The maximal efficiency of brittle failure determined by the energy released in seismic waves (E_q) is also gained in regions with a medium level of stresses. When the level of the deviator and isotropic stresses increases, the $\Delta\tau_n$ value decreases so that the plastic mechanism of relaxation of elastic deformations becomes preferable. Brittle failure becomes impossible in the region to the right of point *A*, i.e., the junction of the limiting envelope and dry static friction line (Fig. 2, point *A*). In [8], Rice noted the greater efficiency of brittle failure in the regions of medium levels of stresses as compared with the regions of low and high levels of stresses.

FAVORABLE CONDITIONS FOR THE ACTIVATION OF DEFECTS

Our investigations demonstrate that a sharp gradient in the distribution of stresses along the activating fault zone is a factor that provides the activation of defects in the medium and creates conditions for more effective release of the energy of elastic deformations during brittle failure. The well-known Griffith's solution, which determines the length of the effects for a given level of stress, which mark the onset of activation, was obtained in the assumption of a homogeneous field of initial stresses. Its extrapolation to a 2D field of initial stresses, which governs the linear variation of stresses along the defect with size $2L$, yields a linear law of variation in the stress drop values:

$$\Delta\tau_n(x) = \bar{\Delta}\tau_n\left(1 + \frac{\eta x}{L}\right) \quad \text{at} \quad |x| \le L \tag{3}$$

This law allows us to write the following expressions for the jump of displacements $\Delta u_x(x)$ on the fracture and shear stress $\tau_n(x)$ formed beyond its limits:

$$\Delta u_x = \frac{(1+\kappa)\bar{\Delta}\tau_n}{2\mu}\left(1 + \frac{0.5\eta x}{L}\right)\sqrt{L^2 - x^2} \quad \text{at} \quad |x| \le L,$$

$$\tau_n = \tau_n^0(x) + \bar{\Delta}\tau_n\left[\frac{x - 0.5\eta(L^2 - 2x^2)/L}{\sqrt{L^2 - x^2}}\operatorname{sgn} x - \frac{\eta x}{L}\right] \quad \text{at} \quad |x| \ge L. \tag{4}$$

In expressions (3) and (4), $0 \le \eta \le 1$ is the coefficient of inhomogeneity of stresses, $\bar{\Delta}\tau_n$ is the mean stress drop for a defect, κ and μ are elastic parameters, and τ_n^0 are initial shear stresses along the *x*-axis. The values of Δu_x and τ_n at $\eta = 0$ correspond to the solution for a fracture in the homogeneous field of initial stresses [7].

It follows from (4) that the ends of the defect are not in equal conditions in the inhomogeneous field of initial stresses, so that Griffith's condition can be gained for one of them (the energy released during its increase exceeds the energy needed to overcome the friction forces and formation of a new surface). At the same time, the other end can remain in the subcritical state. The expression for the energy released during the growth of such a fracture (defect) end over a small length ΔL is written as

$$\Delta W = \frac{(1+\kappa)\bar{\Delta}\tau_n^2}{2\mu}(1 + 0.5\eta\operatorname{sgn} x)^2 L\Delta L. \tag{5}$$

At $\eta = 1$, the energy released during the growth of the activating fracture end is two times greater than in the case of homogeneous distribution of stress drop ($\eta = 0$ is Griffith's solution) and one order of magnitude greater than for a passive end. Thus, the size of the defect, which marks the onset of its transformation into a shear fault, decreases with increasing inhomogeneities in the field of initial stresses. In this case, the fault spreads in the direction of increasing values of stress drop (direction of the gray arrow in Fig. 2), which corresponds to a transition from the level of high deviator and isotropic stresses to the region of medium levels.

It follows from the obtained result that the onset of the catastrophic SAE source opening near the region of high $\Delta\tau_n$ gradients is not occasional. This region is most hazardous from the point of view of the transformation of defects into elongated faults. Thus, zones of a high gradient of stresses should be considered as regions of the metastable state of the Earth's crust [9] that are especially hazardous in terms of the occurrence of strong earthquakes.

HIERARCHY OF MULTISCALE LEVELS OF THE STRESSED STATE

The analysis performed above should be complemented with a note about the hierarchy of the stresses in heterogeneous media. The parameters of the regional stressed state based on the results of tectonophysical reconstruction (averaging over the first tens of kilometers) would seem more inhomogeneous on a smaller averaging scale. The fluctuation of stresses on a subregional scale (averaging over the first kilometers) is performed relative to stresses of the regional averaging scale. A complex mosaic distribution of stresses at higher levels of scales always exists in seismogenic fault zones. These scales should always include regions where brittle faults can develop or activate effectively and plastic flow regions (increased level of isotropic stresses). After the origination in a region with the corresponding limiting state, a brittle fault would necessarily evolve up to a plastic region that would either significantly slow down or stop its development [3].

Formation of the regional fault (a few hundred kilometers) requires the opening of a subregional stress system, where brittle failure is energetically inefficient and accumulation of sufficient energy is needed for the fault during the development of brittle failure. This scenario is possible in the case of the existence of a region of a high gradient of regional stresses. A combination of such a high-gradient region with an elongated region of isotropic stresses of a medium level, where the brittle failure is most efficient, governs the site of preparation of anomalously strong earthquakes.

After the appearance of an elongated region of deviator and isotropic stresses of a medium level, the time of transition to the metastable state, which is sensitive to small energetic perturbations that can provoke a release of enormous energy, would be governed by the rate of the stress gradient formation at the boundary between the regions of medium and high levels of stresses. It is likely that we should discriminate the timing of an elongated region of medium level stresses and the timing of the region with a high stress gradient. It is also possible to suppose that the level of the stress gradient can decrease and a part of the crust can be released from the metastable state due to increase (or decrease) in stresses in the corresponding areas.

CONCLUSIONS

Analysis of regularities in the distribution of natural stresses and their correlation with dynamic parameters of earthquake sources showed the ambiguity of the concept about the preferential origination of a strong earthquake in the regions of high levels of deviator stresses. This concept (Reid's concept), which is the basis of all modern theories of earthquake sources, actually determines only the general idea about the impossibility of long-term existence of such types of regions. The correlation between deviator and isotropic stresses existing in the regions of active seismic regimes shows that relaxation of such high stresses in the respective regions would be related to the quasi-plastic flow due to a large number of weak earthquakes and/or creep flow along the fracture.

Investigation of the distribution pattern of natural stresses in seismoactive regions showed that elongated regions of medium levels of stresses located adjacent to the regions of their high gradient are most hazardous with respect to the possibility of an anomalously strong earthquake. It was found that a high level of the stress gradient provides the realization of smaller defects. The most hazardous area is located at the end of the fault developing in the direction of a high level of stress drop. For rocks located near the limiting state, this state corresponds to a decrease in the level of deviator and isotropic stresses (Fig. 2). Our investigations provide physical grounds of a new paradigm for the prediction of seismic hazard based on the knowledge of regularities in the development of brittle failure of rock massifs.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 06-05-64904 and 07-05-00106.

REFERENCES

1. Yu. L. Rebetskii, Dokl. Earth Sci. **365**, 370 (1999) [Dokl. Akad. Nauk **365**, 392 (1999)].
2. Yu. L. Rebetskii, Geophys. J. **27** (1), 39 (2005).
3. Yu. L. Rebetskii, in *Fluid and Geodynamics* (Nauka, Moscow, 2006), pp. 120–146 [in Russian].
4. Yu. L. Rebetskii and A. V. Marinin, Dokl. Earth Sci. **407**, 321 (2006) [Dokl. Akad. Nauk **407**, 106 (2006)].
5. C. J. Ammon, C. Ji, H.-K. Thio, et al., Science **308**, 1133 (2005).
6. J. D. Byerlee, J. Geophys. Res. **72**, 3639 (1967).
7. D. N. Osokina and V. N. Fridman, *Stress and Deformation Fields in the Earth's Crust* (Nauka, Moscow, 1987), pp. 74–119 [in Russian].
8. 8. J. R. Rice, *The Mechanics of Earthquake Fault*, in *Physics of the Earth's Interior* (Proc. International School of Physics 'Enrico Fermi', Course 78, 1979), Ed. by A. M. Dziewonski and E. Boschi (Italian Physical Society and North-Holland Publ. Co., 1980; Mir, Moscow, 1982).
9. S. V. Gol'din, Izv. Phys. Solid Earth **40**, 817, (2004) [Izv. Ross. Akad. Nauk, Fiz. Zemli, No. 10, 37 (2004)].