ELSEVIER

Earth and Planetary Science Letters 170 (1999) 475–486

EPSL

www.elsevier.com/locate/epsl

Mechanisms of slip nucleation during earthquakes

Ze'ev Reches *

Institute of Earth Sciences, Hebrew University, Jerusalem 91904, Israel

Received 23 December 1998; revised version received 24 May 1999; accepted 24 May 1999

Abstract

Slip nucleation during earthquakes is apparently analogous to rupture nucleation within an intact rock sample subjected to triaxial loading. The observations indicate that both these nucleation processes initiate within a relatively small volume and in both the slip propagates unstably along a quasi-planar surface. In both processes a single, pre-existing, shear fracture cannot nucleate the large-scale slip, and in both a 'process zone' that includes several interacting fractures in a small volume are required to initiate the unstable slip. Both processes require rupture of intact rocks, generate complex fracture geometry, and are associated with intense energy-release rate during slip. Recent observations and analyses are used to correlate rupture nucleation in laboratory tests with nucleation events of large earthquakes. It is proposed that earthquake nucleation occurs by the interaction among multiple fractures within a small volume that develops into unstable yielding of the healed fault zone.

Keywords: earthquakes; propagation; rupture; slip rates; friction; rock mechanics

1. Introduction

Most large earthquakes of the upper crust are slip events along existing fault zones. The nature of this slip has been the subject of many investigations since the last century. Gilbert [1] in his discussion of the 1872 Owens Valley earthquake stated that: "... strain increases until it is sufficient to overcome the starting friction along the fracture." The recognition of 'static' and 'dynamic' friction has remained a fundamental concept of earthquake mechanics. Reid [2], on the other hand, described earthquakes in terms of 'rupture', proposing that "we should expect that the slow accumulation of strain would, in general, reach a maximum value and bring about a rupture in a single, comparatively narrow fault-zone." Recent analyses of strong motion and teleseismic wave data revealed that earthquake slip is characterized by a leading rupture front [3]. Heaton [3] proposed that a 'rupture pulse' moves from the hypocenter area into the locked parts of the fault zone, and he stated that "since ... no slip pulse will propagate unless a slip pulse already exists, the model clearly begs the question of how the rupture pulse starts in the first place" ([3], p. 16). The present work focuses exactly on this question: the mechanism of slip nucleation along a locked fault.

The concepts of *friction* and *rupture* were studied in two different configurations of laboratory models. In the rupture type tests, intact rock samples under confining pressure, σ_3, are subjected to increasing axial loading, σ_1 (Fig. 1A). A typical test displays a few stages during load increase [4]: non-linear ini-

* Tel.: +972-2-658-4686; Fax: +972-2-566-2581; E-mail: reches@earth.es.huji.ac.il

0012-821X/99/$ – see front matter
PII: S0012-821X(99)00122-3

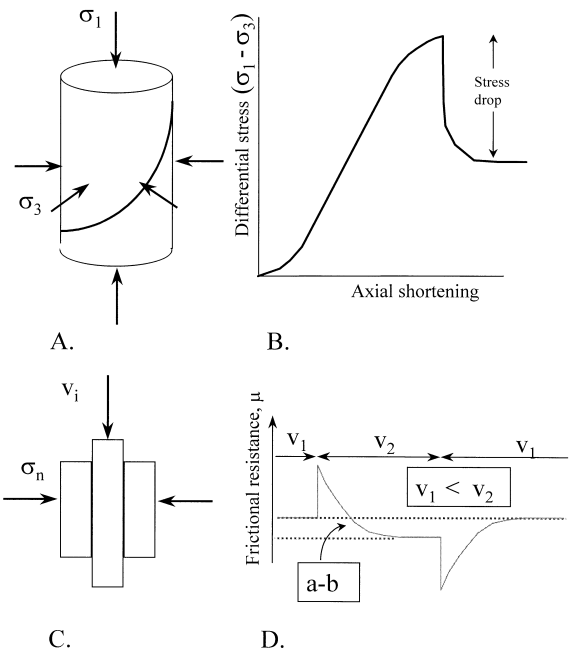

Fig. 1. Laboratory models of earthquake mechanics. (A) An intact rock sample under confining pressure σ_3, in a triaxial configuration with axial loading, σ_1; it typically yields along a quasi-planar surface (dark, curved line). (B) The typical stress–strain relations for the triaxial test with a brittle rock including the post-failure slip (after [4]). (C) One type of friction test: three solid rock blocks subjected to normal stress, σ_n, are enforced to slip at a prescribed constant velocity V_i (other types include saw-cut and rotary samples). (D) Frictional resistance (gray line) in (C) as function of slip velocities V_1 and V_2; velocity weakening is measured by the value of $(a - b)$ (after [5,6]).

tial flaw closure, linear elastic loading, yielding with associated onset of acoustic emission, unstable rupture along a quasi-planar surface at ultimate strength, profound stress drop and stable sliding (Fig. 1B). In the typical friction type test, two (or three) solid rock blocks subjected to normal stress, σ_n, are enforced to slip at a prescribed rate (velocity V_i in Fig. 1C). The friction experiments conveyed the development of rate- and state-variable friction, also known as Dieterich law (Fig. 1D) [5,6]. These models are considered here as idealized end-member configurations for rupture (Fig. 1A) and friction (Fig. 1C); therefore, other testing configurations (e.g. cylinders with saw-cuts or rotary rings) are not discussed.

Some observations of earthquakes are pertinent for the comparison with the laboratory analogs. First, the geometry of fracturing during earthquakes could be highly complex as revealed by surface rupture of recent earthquakes, such as Loma Prieta, 1989 [7,8] and Landers, 1992 [9,10]. Second, earthquakes nucleate within a small volume at the focal region and propagate by 'rupture pulse' along a quasi-planar surface [3]. Third, at a depth of a few kilometers and in the presence of hot water, crushed gouge material is likely to heal and re-cement during periods of interseismic quiescence [11,12]. The healed, re-cemented gouge could behave as an intact rock. Finally, recent analysis of velocity seismograms of several tens of earthquakes revealed that the arrival of the main shock P waves is preceded by a distinct, initial phase [13–15]. This phase was interpreted as representing a few small events that facilitate the nucleation of the main earthquakes. These earthquake features are discussed in detail later.

It will be demonstrated below that earthquake nucleation and slip propagation are similar to rupture nucleation within an intact rock (Fig. 1A,B). The rupture concept that was already proposed in 1910 by Reid was overlooked during the last few decades in favor of the rate- and state-friction law. However, recent observations and analyses of both laboratory rock rupture and earthquake nucleation provide the clear tools to establish this concept. In the following sections we discuss in detail the earthquake features that are in accord with the rupture model and we later present a mechanism for earthquake nucleation based on brittle yielding of intact rocks.

2. Some features of earthquake slip

2.1. Rupture geometry

The 1992, *M* 7.5 Landers earthquake with maximum horizontal displacement of 6 m produced an 80 km long zone of spectacular surface rupture (Fig. 2). Thanks to a combination of excellent exposures and professional interests, the detailed mapping of this event is accurate and illuminating. Johnson et al. [9], who mapped portions of Landers surface, characterized parts of the rupture zone as *belts of shear zones* that are 50–200 m wide with subparallel walls and distributed shear (Fig. 2B). The belts include narrow left- and right-lateral shear zones, and major and minor tensile fractures (Fig. 2B). Slip along individual

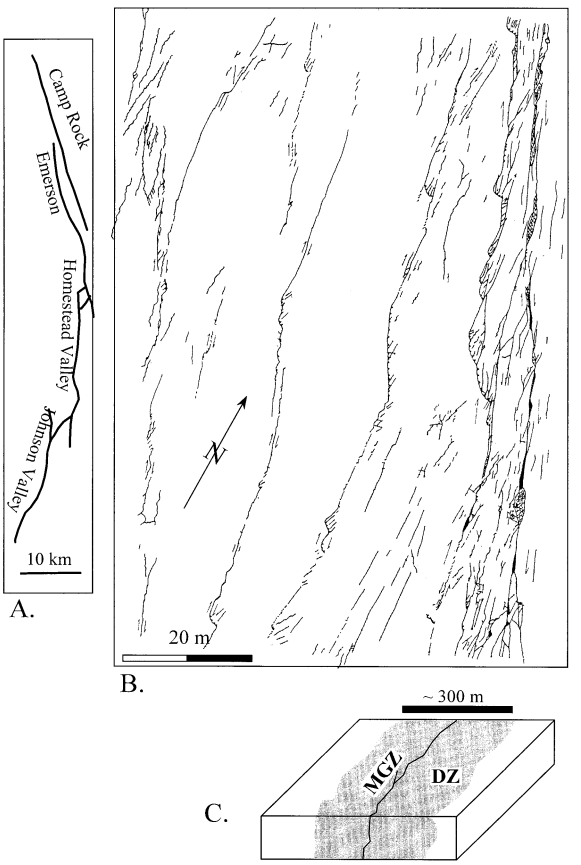

Fig. 2. Features of surface rupture associated with the 1992, *M* 7.5 Landers earthquake. (A) The generalized map of an 80 km long zone of surface rupture; note geometric complexity even at this scale. (B) Detailed map of the surface rupture along part of the Homestead Valley segment; the rupture zone displays a belt of shear zones that is 50–200 m wide with subparallel walls and distributed shear (see text) (after [9]). (C) The generalized geometry of an active fault zone based on the exhumed San Gabriel and Punchbowl faults, California (after [17,18]); *DZ* = damage zone, *MGZ* = main gouge zone.

fractures was a few centimeters in a vertical direction and up to 70 cm of horizontal offset. Ground surface within the shear zones is locally thrust or cut by small grabens. McGill and Rubin [10] presented similar observations along the Emerson fault segment of Landers earthquake (Fig. 2A). They showed that while "the right-lateral slip . . . varied about 150 to 530 cm along the main rupture zone, . . . a total of up to 110 cm of additional right-lateral slip occurred on secondary faults up to 1.7 km away from the main rupture zone."

Is the complex rupture zone of Landers earthquake a surficial feature, or does it reflect the subsurface structure? Johnson et al. [9] claimed that the mapped belts of shear zones as well as individual shear zones are surface expressions of subsurface deformation patterns and structure. Support to this concept was provided by McGarr et al. [16] who analyzed the seismic tremors at 3-km depth in a mine, South Africa. They described a highly fractured, inhomogeneous fault zone and found that "most of the released energy is consumed in creating the fault zone, with less than 1% being radiated seismically". In general, however, the subsurface fracture patterns associated with a *single* earthquake are not accessible. Chester and coworkers [17,18] studied exhumed sections of the North Branch San Gabriel fault and Punchbowl fault, California. These sections carried 16 km or more of right-lateral slip and were buried at depths of 2–5 km. In these sections, the authors recognized 'damage zones' (dotted region marked DZ in Fig. 2C) in which the intensity of deformation increases inward toward a central layer (thin line in Fig. 2C). The damage zone developed during many earthquakes; it is currently a few tens to a few hundreds of meters wide, intensely brecciated and well cemented. Chester and coworkers [17,18] recognized ". . . extreme localization of slip to the central layer, at least during the later stages of fault evolution" ([18], p. 773). It is proposed here, however, that the 'damage zone' (DZ in Fig. 2C) might be the cemented, subsurface equivalent of the 'belts of shear zones' observed as surface rupture during Landers earthquake (Fig. 2B).

It is not clear if the rupture features mentioned above are typical for other large earthquakes. Frequently, surface rupture occurs in loose soil that serves as a 'stresscoat' which obliterates structural details, or on the other hand, could amplify the deformation. Fine details of surface rupture may be eroded in a short period of time, and reliable surface rupture maps can be generated only during days to weeks after the major event. Johnson et al. [19] reviewed several previous major earthquakes, including 1906 San Francisco, 1972 Managua, and 1968 Borrego Mountain, and concluded that Landers structural features are similar to structures developed during comparable events. Yet, one can also expect that while many new fractures formed during an

earthquake, most of the slip of this event was concentrated along a narrow slip surface (e.g. [17,18]).

The above descriptions of the intricate shear belt in Landers and other faults have far-reaching implications. The most important one is that a large earthquake generates a damage belt that is tens to hundreds of meters wide. Within the belt, the shear is accommodated by a combination of tensile fractures, thrusts, narrow shear zones, distributed deformation, as well as localized slip along the main fault surface. This configuration implies that even earthquakes along mature faults could generate networks of many, new fractures.

2.2. Healing of a fault zone

The crushed and cracked rock fragments within a fault zone are likely to heal relatively fast under the thermal conditions prevalent at earthquake nucleation depth. Bodnar and Sterner [20] examined healing in quartz. They noted that "... at temperatures of 400°–500° the fractures begin to heal almost immediately", and almost complete healing was achieved after hours to days above 400°C.

Karner et al. [11] studied fault gouge healing under hydrothermal conditions. They ran triaxial tests on saw-cut samples of Sioux quartzite with fine grain quartz gouge. The samples were deformed under a confining pressure of 250 MPa, temperature of 230°–636°C, and pore fluid pressure of 75 MPa. The tests included heal–slide or slide–heal–slide history with healing times ranging from 1 h to 28 h. A representative experiment of healing at 636°C followed by sliding at 230°C is shown in Fig. 3 together with an unhealed sample deformed at 230°C. The healed gouge displays a clear and significant strength increase. In the heal–slide experiments, the authors found that the gouge friction coefficient increases by 0.1 per decade of seconds of healing time without apparent saturation. This result indicates that during healing times of months and years, the gouge could regain the strength of the host quartzite.

In-situ conditions at the nucleation depth of large earthquakes along the San Andreas, for example, generally correspond to the conditions in the healing experiments cited above. The earthquakes nucleate at depths of 7–15 km, where the temperature is approximately 250°–400°C, and the rocks are saturated

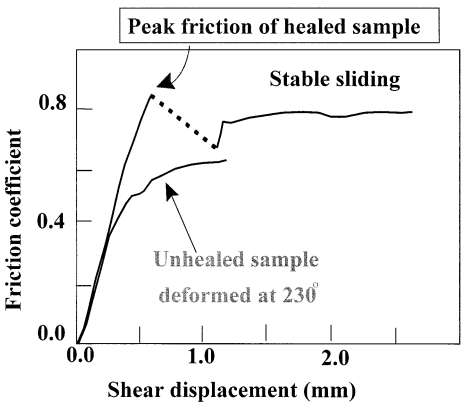

Fig. 3. Frictional strength of cylindrical samples (Fig. 1A) of Sioux quartzite with a saw-cut filled with fine grain quartz gouge [11]. The confining pressure was 250 MPa and pore fluid pressure of 75 MPa. Two tests are shown: (1) healing at 636°C followed by sliding at 230°C (thin line); and (2) an unhealed sample deformed at 230°C (thick line). Note that the healed gouge is significantly stronger than the unhealed sample (friction coefficient $\mu > 0.8$ at peak strength); yielding is followed by a large friction drop of $\Delta\mu \approx 0.17$. The unhealed sample slips stably.

with water or brines. Each earthquake causes intensive crushing that leaves many shattered grains and a fault zone under nonuniform high stresses. The high temperature and the presence of pore water would facilitate local dissolution and local precipitation of the crushed gouge driven by pressure-solution processes. These processes are evident, for example, in abundance of veins with secondary mineralization [17,21]. The interseismic periods for great earthquakes in California are 50–300 years, and it is reasonable to assume that healing would lead to partial or complete recovery of the fault zone. The earthquake rupture encounters a re-cemented fault zone that is not loose gouge or smooth cohesionless surface.

2.3. Fracture energy

The energy-release rate G is probably the best parameter to evaluate the 'difficulty' to form and propagate a fracture [22,23]. It is defined as the energy flux to the fracture tip per unit length of fracture growth per unit width along the fracture front; namely, G is equivalent to the energy per unit area of the fracture. The parameter G integrates the energy

Table 1
The energy-release-rate of slip along shear discontinuities after Li [22]

Rock type or area	Slip type	G (J m^{-2})	Reference
Smooth granite surface	Low normal stress	0.1–2.5	[24]
Carbonates and sandstones	Joint surfaces	10–1000	[22]
Granite and gabbro	Intact, triaxial	$(0.3–5) \times 10^4$	[22]
Creep zone, California	Geodetic data	$(6–30) \times 10^6$	[22]
Earthquakes, general	General	10^6–10^8	[22]

consumed by all the mechanisms associated with fracturing: seismic radiation, heat production, plastic deformation, and generation of new surface area and acceleration. Li [22] showed various methods to calculate G that are based on fracture geometry, stress and strain intensities and elastic parameters. He calculated and compiled G values for triaxial failure of intact rocks, slip along saw-cut samples and joints, slip along natural faults and during earthquakes. A few typical values appear in Table 1 and Fig. 4.

The experimental healing rates (Fig. 3) suggest that fault gouge re-cements, partly or completely, during a few years to tens of years. Therefore, most, if not all, fractures that form during an earthquake should be regarded as new fractures. The energy values of Table 1 can be used to estimate the amount of this new fracturing. Scholz ([25], p. 167) mentioned that G indicates substantial volume of inelastic deformation at the earthquake propagating tip.

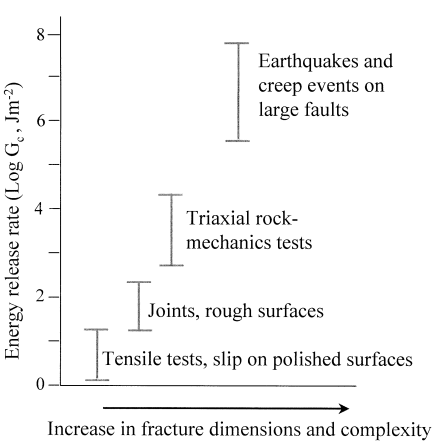

Fig. 4. The energy-release rate determined for triaxial failure of intact rocks, slip along saw-cut samples and joints, slip along natural faults and during earthquakes (after [22]).

In addition to plastic flow, this inelastic deformation also includes new surface area associated with fracturing. We first assume that 1% of the energy-release rate G is consumed by new surface area at the earthquake propagating tip. For $G = 10^6$–10^8 J m^{-2} (Table 1) the energy consumed by new surface area is $E_s = 0.01\ G = 10^4$–10^6 J m^{-2}. The intrinsic surface energy of silicate minerals is $\mu = 1$–10 J m^{-2} [26]; therefore, the new surface area generated during an earthquake is $S = (E_s/\mu) = 10^3$–10^6 m^2 for each m^2 of fault area. In the field, this new surface area is manifested in crushing grains of gouge zones [27,28], and in new fracture zones [9,16].

The assumption that only 1% of G is consumed by new surface area is a conservative one [16]. Boler [29] tested the tensile fracturing of glass plates and found that the energy of radiating elastic waves is smaller than 0.001 of the energy associated with new areas. Olgaard and Brace [30], who studied an active fault in a mine concluded that "... new surface produced and, therefore, the surface energy associated with seismic faulting in the mines, suggests that this energy could contribute more to the total energy budget than has been previously calculated."

2.4. Summary

Several central observations of earthquakes were outlined above. First, the surface rupture of earthquakes displays highly complex patterns of fractures (Fig. 2). Second, the high potential for gouge healing at earthquake nucleation depth (Fig. 3) most likely causes re-cementation of fault gouge during interseismic quiescence periods. Third, the intense energy-release rate associated with earthquakes (Fig. 4) reflects intense energy consumption at the earthquake propagating front.

These features apparently do not fit the friction tests (Fig. 1C,D). In these tests the fault surface does not have physical fault tips and it is not bounded by physical termination zones (Fig. 1C). The slip along such a fault occurs without the resistance at the fault tip, and the energy consumption associated with propagation is neglected. Further, the blocks are not cemented to each other and the energy required for crushing a fault zone is ignored. Finally, the values of rate- and state-dependent friction are based on slip at constant velocity, but the initiation and propagation of slip are not considered.

On the other hand, the features of fault zone complexity, fault zone cementation and high energy-release rate are similar to features of intact rock rupture. To demonstrate this similarity, I outline in the next section the experimental observations of slip nucleation in intact rocks, and later, these observations are related to earthquake nucleation.

3. Slip nucleation in laboratory tests

3.1. Experimental observations

It was suggested above that slip nucleation and propagation during earthquakes would be better described in terms of slip nucleation during rock rupture. Slip nucleation within a solid, brittle sample was experimentally explored with three different configurations described below. The first two are based on the assumption that a weak planar surface localizes the shear and initiates macroscopic slip; the third configuration assumes that slip nucleates spontaneously within a solid rock.

3.1.1. A single pre-loading fracture

Many investigators attempted to verify Griffith's model that a single, critically oriented fracture under shear could propagate as a fault. The most comprehensive analysis of this topic was recently presented by Germanovich et al. [31]. They developed perfect penny-shaped, 3-D, fractures by focusing a laser beam inside a transparent sample of plastic or glass. This method allows to control size, orientation and location of the initial fractures. The striking results in both 2-D and 3-D cases is that a pre-cut fracture under shear stress does not grow in its own planes and does not propagate in the direction of the maximum shear stress. Rather, tensile wing-cracks develop at the tips of the pre-cut fracture and these cracks propagate while curving toward the axis of maximum compression. Wing-crack propagation is stable, in the sense that incremental crack propagation requires incremental increase of axial load. Further, the 3-D wing cracks of Germanovich et al. [31] could not propagate unboundedly to break the sample, and the authors concluded that "Unlike 2-D cracking, there are intrinsic limits on 3-D growth of wing cracks produced by a single preexisting crack."

3.1.2. An array of pre-loading fractures

The unsuccessful attempts to nucleate slip from a single fracture, led to experimental attempts with arrays of well-organized fractures inserted prior to loading [31,32]. These experiments succeeded in producing a through-going fault that developed by linkage between well-organized fractures. Therefore, the experiments with pre-cut fractures portray that in brittle materials, a single, pre-cut fault does not grow by in-plane shear propagation even for the perfect fracture orientation. Further, "large (macro) crack propagation can result *only* from the combined action of wing cracks originating from several properly situated initial cracks" [31].

3.1.3. An intact, brittle rock without pre-cuts

If a small fracture cannot serve as the nucleus for a large fault, then how does slip nucleate in intact, brittle rocks? Lockner et al. [33] used locations of acoustic emission events (AE) within a triaxially loaded granite sample to monitor fault nucleation. These experiments provide a few key observations related to the above question: (1) The fault nucleates within a *small* zone without a pre-faulting alignment of AE events. In Lockner's experiments, the events associated with the nucleation are restricted to a region of about 1–3 cm^3 in volume. (2) The nucleation stage in the granite experiments appears to display two phases. During the first one, few AE events occur within a small region without clear pattern. This is followed by a second phase during which the AE events are organized along a *planar* surface. The second phase indicates the initial growth of the fault. (3) After the nucleation, the fault propagates from the nucleation zone by in-plane growth as mixed mode

II and mode III fracture. (4) The nucleation and propagation of the fault are characterized by intense local cracking, occurring within a process zone that leads the propagating fault.

3.2. A model for slip nucleation within an intact rock: tensile cracks interactions

Reches and Lockner [34] used the observations of Lockner et al. [33] to develop a fault nucleation model that is based on one assumption: faults nucleate and grow by interaction among dilating cracks that are parallel to maximum compressive stress (Fig. 5). The model considers a thin brittle plate that contains many potential cracks and which is biaxially loaded. During early loading some cracks dilate and alter the stress field in their proximity, but because the dilating cracks are sparse and randomly distributed, they do not affect each other. The density of the dilated cracks increases nonlinearly with increasing load, and locally a crack that otherwise would be stable yields due to the tensile stress induced by its dilated neighbor. In the highly stressed plate, the dilation of one crack can induce the dilation of closely spaced neighboring cracks that were themselves on the verge of opening. This region with dense interacting cracks, called the *process zone*, extends itself by inducing tensile stresses across new cracks. As the process zone lengthens, its central part is weakened due to the dense cracking and eventually yields by shear. This locus of sheared, rotated and crushed blocks is the fault nucleus that develops into the fault zone.

Reches and Lockner [34] calculated the stress fields associated with tensile and shear cracks to evaluate the intensity of dilation interaction of cracks. Clearly, interaction intensity depends on the relative position of the new crack with respect to the first, dilating one. The new crack, which is most likely to dilate due to stress induction of the first

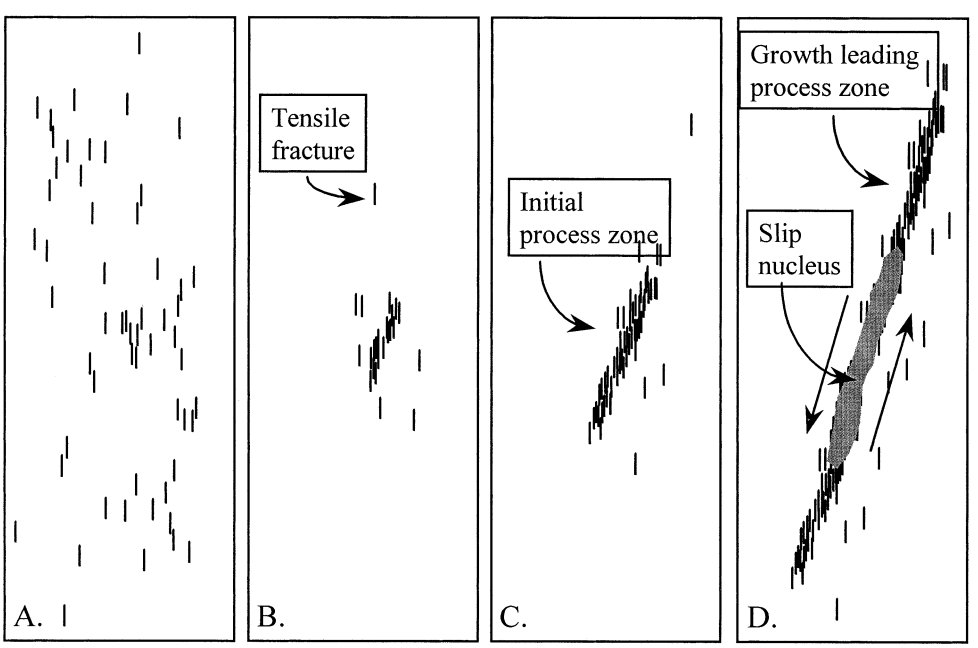

Fig. 5. The nucleation and growth of a fault in a brittle rock by interaction among dilating cracks that are parallel to maximum compressive stress (after [34]). (A) Initial stage in which a brittle rock plate contains many potential cracks; during early axial loading some cracks dilate but they do not interact with each other. (B) The density of the dilated cracks increases with increasing load, and locally a crack that otherwise would be stable yields due to the tensile stress induced by its dilated neighbor. (C) The dilation of one crack can induce the dilation of neighboring cracks that were themselves on the verge of opening; the region of interacting cracks is called the process zone. (D) The process zone extends itself by inducing tensile stresses across new cracks and lengthens; its central part is weakened due to the dense cracking and eventually yields by shear; this locus of sheared, rotated and crushed blocks is the fault nucleus that develops into the fault zone.

crack, is *stepped* in a critical position (Fig. 5). The recognition that an induced crack is off the axis of the first crack is central to the model, because the orientation of the eventual shear fracture is controlled by this off-axis constructive crack interaction [32,35]. Further, one can show that an array of two, three or more interacting cracks, stepped in the critical geometry, would form a process zone of subparallel tensile fractures (Fig. 5). Due to the increase of the induced tensile stress the process zone extends itself unstably, inducing tensile cracking, and generates the *fault nucleus* [34]. This process is termed *self-organized-cracking*.

The self-organized-cracking model that is based on one assumption (faults nucleate by interaction of dilating cracks) leads to four main predictions [34]. (A) Interacting cracks are stepped with respect to each other and diagonal to the fault zone that develops. (B) The tension induced by the dilating and interacting cracks generates a nucleation zone that propagates in a self-organized manner, maintaining its orientation with respect to the loading stresses; the fault propagates in its own plane. (C) Shear nucleation and growth by crack interaction is an unstable process due to enhancement of the induced tensile stresses. (D) The model further predicts that the angle between the fault and the axis of maximum compression is 20º to 30º. These results are in agreement with the experimental observations of fault nucleation and growth outlined above. In the following section we discuss the applicability of the self-organized-cracking model to earthquake nucleation.

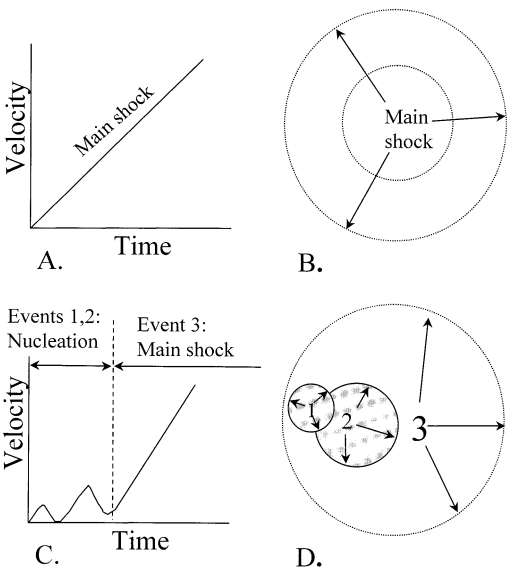

Fig. 6. A nucleation stage model for earthquakes (after [14]). (A) The linear velocity increase that is anticipated for an earthquake which obeys the scale-independent behavior. (B) The anticipated slip distribution for the main shock of (A). (C) Schematic presentation of the velocity seismogram observed by [14,15] in a few tens of earthquakes. The main shock of P waves is preceded by a distinct, initial phase. (D) The slip history proposed by Ellsworth and Beroza ([14] for (C); two small events, *1* and *2*, assumed to precede the main event, *3*.

4. Discussion

4.1. Seismic evidence of slip nucleation during earthquakes

An earthquake nucleation stage was detected recently by careful analysis of velocity seismograms of large, shallow earthquakes [14,15], as well as microearthquakes [13]. In general, a linear velocity increase is anticipated for an earthquake that obeys the scale-independent behavior [15]. The corresponding velocity seismogram should appear as in Fig. 6A for the scheme of slip propagation shown in Fig. 6B. However, Umeda [14] and Ellsworth and Beroza [15] found in the records of few tens of earthquakes that the arrival of the main shock of P waves is preceded by a distinct, initial phase (Fig. 6C). This phase is characterized by a low rate of seismic moment release and low velocity that is followed by the sudden increase of the seismogram velocity. The duration and moment of this phase are related to the magnitude of the eventual earthquake, thus contradicting the scale-independent relations of earthquakes. Iio [13] recognized a ‘slow initial phase’ in 69 microearthquakes that he analyzed from the aftershock set of Western Nagano Prefecture earthquake, 1984, Japan. He demonstrated that the initial phase is not an instrumental effect or the result of basement structure, and concluded that this phase reflects the nucleation stage of the microearthquakes. In the above-mentioned works [13–15] the initial phase appeared in all analyzed earthquakes that range in magnitude from −0.7 to 8.1 (moment magnitude).

Ellsworth and Beroza [15] suggested that this

Table 2
Comparisons of the nucleation phases in experiments [33], models [34] and earthquakes [15]

Nucleation phases	AE events in laboratory experiments [33]	Self-organized-cracking [34]	Nucleation and cascade model [15]
Phase I: pre-nucleation, low global stresses	Hundreds of AE events in random distribution; high *b*-values (relative abundance of weak AE); rising stresses.	Tensile cracking in random distribution (strength variations); no interactions between cracks.	Widespread, random seismic activity along the fault-zone (weak, cemented zone).
Phase II: early nucleation, cascading stage	Tens of AE events within a small volume; development of a weak region that is the initial process zone; no fault pattern; no stress drop.	Localized tensile cracking associated with intense interaction among close, stepped cracks; high stresses; process zone forms.	A few, small shear events that lead to mutual delayed failure; slow growth; low seismic moment; slow increase of seismic velocity.
Phase III: nucleation, formation of a planar surface	The 20–30 AE events after phase II form a planar, quasi-circular process zone; negligible stress drop.	The intense stress interaction among the stepped cracks forms a planar process zone.	The few earlier events form the main shock nucleus; radius and moment of nucleus are proportional to magnitude of main shock.
Phase IV: shear yielding, unstable shear propagation	Thousands of AE events occur within a narrow process zone, along the front of the propagating fault; large stress drop.	The process zone induces tensile stress and dilates cracks along its front; it propagates unstably; shear yielding in the weak fault nucleus; large stress drop.	Propagation of main earthquake as a 'slip-pulse' along the fault zone.

initial phase corresponds to the nucleation stage of the earthquakes and proposed two conceptual models for this phase. One is the *cascade* model with a small event that leads to a delayed failure of a larger, second event; the two events together eventually lead to the main shock (Fig. 6D). The second model is the *pre-slip* model in which aseismic slip precedes a few small events confined to the pre-slip area, and this activity leads to the main shock. The authors noted that the two models could not be distinguished due to the limited resolution of seismic data.

The nucleation phases recognized in the seismic record [13–15] can be correlated with the experimental observations of slip nucleation [33], and the associated model of Reches and Lockner [34] (Fig. 5). Figs. 5 and 6 and Table 2 display the suggested relations of the different models. The most critical similarity between the experimental nucleation and the earthquake nucleation is the occurrence of a few fractures close to each other during phases II and III (Table 2). It is claimed that *interaction among several fractures* is a necessary condition for slip to nucleate. This is in agreement with the well-known observations that a single, shear flaw cannot grow [31]. Once several fractures interact, they form the initial process zone, and the slip propagates unstably as either an earthquake along the cemented fault zone, or as a new fault zone that grows within an intact rock.

4.2. Geometry of earthquake nucleation zone

While the concept of nucleation is in general agreement in the cited studies (Table 2), the geometry of the fractures that lead to failure differs. This geometry may not be ignored as it could control fault properties. For example, the angular relationships within an array of en-echelon, stepped fractures control the intensity of stress interaction and thus the intensity of slip instability [32,34,35]. Or, the occurrence of slip along secondary Riedel shears within a fault zone apparently controls the strength of the fault [36]. Merzer and Klemperer [37] presented another interesting case. They claimed that pre-earthquake linkage of fault-parallel cracks could explain the anomaly of electromagnetic radiation at the ultra-low frequency range that was observed before the 1989 Loma Prieta earthquake. In this section

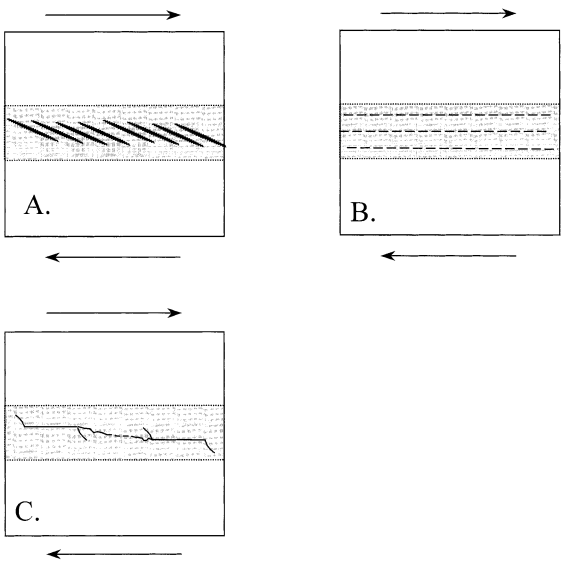

Fig. 7. Fracture patterns that were proposed for slip nucleation in earthquakes. (A) En-echelon tensile fractures arranged in a self-organized pattern (after [34]). (B) Fault-parallel shear fractures (after [38]). (C) Pairs of shear fractures (after [39,40]).

we describe and discuss some fracture patterns proposed for slip nucleation.

4.2.1. En-echelon tensile fractures

Fig. 7A displays the model in which earthquake nucleation occurs due to interaction of tensile fractures arranged in an en-echelon pattern [34]. This pattern leads to enhanced stress induction on the interaction cracks and thus to unstable propagation (see above and [34]).

4.2.2. Fault-parallel shear fractures

Fig. 7B displays the model of Kame and Yamashita [38] who simulated the dynamic interaction and linking of many cracks that are initially equal in length, and parallel to each other and to the fault zone (Fig. 7B). The authors compared cases of collinear cracks and stepped cracks and suggested that a collinear array of disconnected cracks will link to each other in a gradual growth. The large surface that will eventually form is the earthquake nucleus.

4.2.3. Pairs of shear fractures

Fig. 7C displays the nucleation concept of Trifu and Urbancic [39]. In a detailed study of spatial relationships during the fracturing history associated with a rockburst of magnitude 2.9 in a mine, they attributed the main event to the interaction between two smaller fractures, which were properly oriented with respect to each other. They concluded that the configuration of collinear shear fractures proposed by Shen et al. [40] (Fig. 7C), best fits their observations.

Selecting the 'correct' fracture pattern in the nucleation zone is not an easy task. The limitations of seismic data restrict the resolution of the nucleation stage in terms of individual fractures [15]. Further, the scaling from small size samples to field size faults is unknown, and apparently the concept of self-similarity cannot be employed as this concept is probably invalid with regard to earthquake nucleation [15,41]. These limitations indicate that at the current time, fracture patterns in the nucleation site of an earthquake could be derived by quantitative evaluation of Fig. 7 configurations and not by direct seismic observations.

5. Conclusions

The nucleation of earthquakes and the associated unstable slip propagation should be analyzed in terms of strength, unstable yielding and rupture of intact rocks. This conclusion is based on observations associated with earthquakes: (1) complexity of the fracture systems in the fault zone; (2) the partial or complete healing of crushed fault gouge at great depth; (3) the high values of energy-release rate during earthquakes. Models of slip nucleation in triaxial experiments and seismic data appear to have similar properties. It is recognized that a single shear fracture cannot nucleate an earthquake, and that few interacting fractures within a nucleus are required to initiate the unstable slip. The models differ in details of fracture patterns and mechanisms. It is concluded that earthquakes nucleate due to fracture interaction within a small volume in an intact fault zone, but the detailed pattern of the nucleation fractures and their scaling are still unknown.

Acknowledgements

The concepts presented in this work were developed during discussions with Dave Lockner, Yehuda

Ben-Zion, Jim Dieterich, Zvi Garfunkel, Tom Dewers, Amotz Agnon and Gene Scott. None of the above necessarily agrees with these ideas. Liza Heller-Kalai helped in style editing. The critical reviews of Steve Karner and an anonymous reviewer significantly improved the paper. The research was supported in part by the Rock Mechanics Institute, Oklahoma University. ***[RV]***

References

[1] G.K. Gilbert, A theory of the earthquakes of the Great Basin, with a practical application, Am. J. Sci. 27 (1883) 49–53.
[2] H.F. Reid, The California earthquake of April 18, 1906; the mechanics of the earthquake, Vol. II, Report State Earthquake Investigation Commission, 1910, 192 pp.
[3] T.H. Heaton, Evidence for and implications of self-healing pulses of slip in earthquake rupture, Phys. Earth Planet. Inter. 64 (1990) 1–20.
[4] W. Waswersik, W.F. Brace, Post-failure behavior of a granite and a diabase, Rock Mech. 3 (1971) 61–85.
[5] J.H. Dieterich, Time dependent friction in rocks, J. Geophys. Res. B 77 (1972) 3690–3697.
[6] C.H. Scholz, Earthquakes and friction laws, Nature 391 (1998) 37–42.
[7] A. Aydin, A.M. Johnson, R.W. Fleming, Right-lateral-reverse surface rupture along the San Andreas and Sargent faults associated with the October 17, 1989, Loma Prieta, California, earthquake, Geology 20 (1992) 1063–1067.
[8] A.M. Johnson, R.W. Fleming, Formation of left-lateral fractures within the Summit Ridge shear zone, 1989 Loma Prieta California, earthquake, J. Geophys. Res. B 98 (1993) 21823–21837.
[9] A.M. Johnson, R.W. Fleming, K.M. Cruikshank, Shear zones formed along long, straight traces of fault zones during the 28 June 1992 Landers California, earthquake, Bull. Seismol. Soc. Am. 84 (1994) 499–510.
[10] S.F. McGill, C.M. Rubin, Surficial slip distribution on the central Emerson fault during the June 28, 1992 Landers earthquake, California, J. Geophys. Res. B 104 (1999) 4811–4833.
[11] S.L. Karner, C. Marone, B. Evans, Laboratory study of fault healing and lithification in simulated fault gouge under hydrothermal conditions, Tectonophysics 277 (1997) 41–55.
[12] C. Marone, The effect of loading rate on static friction and the rate of fault healing during the earthquake cycle, Nature 391 (1998) 69–72.
[13] Y. Iio, Observations of the slow initial phase generated by microearthquakes — implications for earthquake nucleation and propagation, J. Geophys. Res. B 100 (1995) 15333–15349.
[14] Y. Umeda, The bright spot of an earthquake, Tectonophysics 211 (1990) 13–22.
[15] W.L. Ellsworth, G.C. Beroza, Seismic evidence for an earthquake nucleation phase, Science 268 (1995) 851–855.
[16] A. McGarr, S.M. Spottiswoode, N.C. Gay, W.D. Ortlepp, Observations relevant to seismic driving stress, stress drop, and efficiency, J. Geophys. Res. B 84 (1979) 2251–2261.
[17] F.M. Chester, J.M. Logan, Implications for mechanical properties of brittle faults from observations of the Punchbowl fault zone, California, Pure Appl. Geophys. 124 (1986) 79–106.
[18] F.M. Chester, J.P. Evans, R.L. Biegel, Internal structure and weakening mechanisms of the San Andreas Fault, J. Geophys. Res. B 98 (1993) 771–786.
[19] A.M. Johnson, K.M. Cruikshank, R.W. Fleming, Borrego Mountain, Loma Prieta, Landers, Northridge; simply earthquakes or seismostructural events? EOS 75 (Suppl. 343) (1994).
[20] R.G. Bodnar, S.M. Sterner, Synthetic fluid inclusions, in: G.C. Ulmer, H.L. Barnes (Eds.), Hydrothermal Experimental Techniques, Wiley, New York, 1987, pp. 423–457.
[21] R.H. Sibson, Earthquake rupturing as a mineralizing agent in hydrothermal systems, Geology 15 (1987) 701–704.
[22] V.C. Li, Mechanics of shear rupture applied to earthquake zones, in: B.K. Atkinson (Ed.), Fracture Mechanics of Rocks, Academic Press, New York, 1987, pp. 351–428.
[23] B. Lawn, Fracture of Brittle Solids (2nd ed.), Cambridge Univ. Press, New York, 1993, 378 pp.
[24] P.G. Okubo, J.H. Dieterich, State variable fault constitutive relations for dynamic slip, in: S. Das et al. (Eds.), Earthquake Source Mechanics, AGU Geophys. Monogr. 37 (1986) 25–35.
[25] C.H. Scholz, The Mechanics of Earthquakes and Faulting, Cambridge Univ. Press, New York, 1990, 439 pp.
[26] M. Friedman, J. Handin, G. Alani, Fracture-surface energy of rocks, Int. J. Rock Mech. Min. Sci. 9 (1972) 757–766.
[27] E.H. Rutter, R.H. Maddock, S.H. Hall, S.H. White, Comparative microstructures of natural and experimentally produced clay-bearing fault gouges, Pure Appl. Geophys. 124 (1986) 3–30.
[28] C. Sammis, G. King, R. Biegel, Kinematics of gouge deformation, Pure Appl. Geophys. 125 (1987) 777–812.
[29] F.M. Boler, Measurements of radiated elastic wave energy from dynamic tensile cracks, J. Geophys. Res. 95 (1990) 2593–2607.
[30] D.L. Olgaard, W.F. Brace, The microstructure of gouge from mining-induced faults near Johannesburg, South Africa, EOS 62 (1981) 400.
[31] L.N. Germanovich, L.M. Ring, B.J. Carter, A.R. Ingraffea, A.V. Dyskin, K.B. Ustinov, Simulation of crack growth and interaction in compression, in: Fujii Toshio (Ed.), Proc. Congr. Int. Soc. Rock Mech. 8, 1995, pp. 219–226.
[32] H. Horii, S. Nemat-Nasser, Compression-induced microcrack growth in brittle solids; axial splitting and shear failure, J. Geophys. Res. B 90 (1985) 3105–3125.
[33] D.A. Lockner, J.D. Byerlee, V.S. Kuksenko, A.V. Pono-

marev, A. Sidorin, Quasi-static fault growth and shear fracture energy in granite, Nature 350 (1991) 39–42.
[34] Z. Reches, D.A. Lockner, The nucleation and growth of faults in brittle rocks, J. Geophys. Res. B 99 (1994) 18159–18174.
[35] Y. Du, A. Aydin, Interaction of multiple cracks and formation of echelon crack arrays, Int. J. Numer. Anal. Methods Geomech. 15 (1991) 205–218.
[36] D.E. Moore, D.A. Lockner, The role of microcracking in shear-fracture propagation in granite, J. Struct. Geol. 17 (1995) 95–114.
[37] A.M. Merzer, S.L. Klemperer, High electrical conductivity in a model lower crust with unconnected, conductive, seismically reflective layers, Geophys. J. Int. 108 (1992) 895–905.
[38] N. Kame, T. Yamashita, Dynamic nucleation process of shallow earthquake faulting in a fault zone, Geophys. J. Int. 128 (1997) 204–216.
[39] I. Trifu-Cezar, T.I. Urbancic, Fracture coalescence as a mechanism for earthquakes; observations based on mining induced microseismicity, Tectonophysics 261 (1997) 193–207.
[40] B. Shen, O. Stephansson, H.H. Einstein, B. Ghahreman, Coalescence of fractures under shear stresses in experiments, J. Geophys. Res. B 100 (1995) 5975–5990.
[41] Y. Umeda, T. Yamashita, T. Tada, N. Kame, Possible mechanisms of dynamic nucleation and arresting of shallow earthquake faulting, Tectonophysics 261 (1996) 179–192.