Growth of geologic fractures into large-strain populations: review of nomenclature, subcritical crack growth, and some implications for rock engineering

R.A. Schultz*

Geomechanics-Rock Fracture Group, Geological Engineering Program, Department of Geological Sciences/172, Mackay School of Mines, University of Nevada, Reno, NV 89557, USA

Accepted 7 October 1999

Abstract

Several aspects of fracture arrays are reviewed briefly and discussed. The terminology applied to progressive or multi-stage brittle deformation in rock masses is improved by noting fundamental mechanical differences in fracture type and the kinematic coupling between dilatant mixed-mode crack displacements and wing cracks developed at the fracture tips. An array of initially mixed-mode (I–II) cracks will evolve under remote tensile least principal stress and with increasing strain to a dilatant, mode-I crack array oriented approximately perpendicular to the remote tensile stress. This progressive fracture growth thus defeats predictions of fracture-set orientation and displacement based only on a Mohr circle estimate of initial elastic stress (valid in the rock mass only at the earliest stages of fracture nucleation). Slow, subcritical crack growth in rock is associated with distinctive changes in fracture population geometry, as shown by published numerical simulations of fracture–network evolution. An increase in the stress corrosion index promotes joint clustering and significant changes in joint length–frequency that may lead to characteristic differences in the statistics of large-strain fracture populations. These geometric clues can be used to refine estimates of strength and deformability of rock masses and to infer classes of physico-chemical processes acting at the fracture tips during the development of the fracture population. © 2000 Elsevier Science Ltd. All rights reserved.

1. Introduction

An understanding of the patterns, characteristics, and genesis of fracture sets in rock is an important element of rock engineering and rock mechanics. The geometric description of rock-mass structure, based on recognition of the strength-reducing effects of fractures (collectively termed “discontinuities” in rock engineering) [1–3] is integral to rock mass classification systems and associated project design [4–7]. However, considerable insight can be gained from a detailed examination of the fractures themselves beyond this initial statistical sampling [8,9], leading in many cases to a more physically based model of rock mass behavior [10].

Advances in the mechanics of fracture development [11–13] have been paralleled by related yet largely independent work on the statistical physics of fracture systems [14–19]. The key to both areas is the nature of stress concentration at the fracture tip, given that the processes of fracture propagation [20], and interaction and linkage [21–24], are either controlled by or related to this small region. As a result, informative correlations may be drawn between the processes acting at the scale and location of the fracture tips and the aggregate properties of fracture sets that are so apparent and important at the outcrop scale.

In this paper, the role of stress–based calculations in predicting fractures in rock and the associated plethora of fracture nomenclature is discussed. Historically this
array of terms stems from the mismatch between a comparison of infinitesimal elastic stress with finite geologic fracture strains. Slow, subcritical crack growth is then emphasized as an important process for regulating fracture-population characteristics and statistics. These results all reinforce an ongoing shift toward post-yield, elastic–plastic formulations of fractures in rock and away from standard engineering-based linear–elastic fracture mechanics (LEFM) approaches.

2. Peak-strength criteria and prediction of fracture type

Considerable discrepancies exist in the literature over the application of Mohr-circle estimates of stress and the associated structures to be expected in the outcrop (see, e.g., Price and Cosgrove [25] and Schultz and Zuber [26] for discussion). Several sources of error in interpretation can lead to incorrect relationships between stress state and the associated brittle structure.

1. Resolution of remote stress state onto a fracture surface suggests only the next small increment of displacement. Infinitesimal elastic stresses by definition cannot be linearly related to large-strain fracture sets [11], making a direct prediction of brittle deformation from elastic models [27] lacking in physical basis unless corrections for propagation, interaction, and linkage processes can be made.

2. In a Mohr circle analysis the magnitudes of stress must be matched to the level that can be sustained by the rock mass prior to failure [26]. Physically this level represents the state of stress near the onset of dilatancy and strain softening in the rock mass [15]; continued strain beyond the elastic limit is typically associated with nucleation and interaction of microcracks and the beginnings of macrocrack linkage and coalescence as the peak strength of the rock mass is reached. Continued strain (for displacement-controlled remote loading) can lead to production of fracture networks [11,28] and faulted joints [29–32] in a test sample or outcrop.

3. Idealization of oblique opening of a fracture, due to tensile and nonzero shear resolved stresses, using an indefinitely long fracture — as implied in a Mohr based analysis — can lead to an equivocal interpretation of the resulting structures. By noting the development of end cracks (see below) at the tips of discontinuous [12,33] mixed-mode (I–II) cracks a clearer view of the fracture kinematics [13] can be achieved (discussed below).

2.1. Kinematic fracture classification

The most useful and robust approach to identifying and classifying fractures in rock involves a definition of their strain significance by noting the sense of displacement (normal or parallel) relative to the fracture plane [12,34]. The resulting classification (e.g., Fig. 1) yields the relevant mechanical interpretation with a minimum of assumptions and speculations about remote stress, geologic history, and other interpretive details. Pollard and Segall [21] outline this approach and its principal assumptions. “Fracture” is used in this paper as a general term for an approximately planar break in rock that identifies a localized discontinuity in displacement at some scale of observation [35]. In this mechanical context all types of cracks, faults, slip surfaces, anticracks, and deformation bands are considered to be stress concentrating, and stress redistributing, fractures.

Displacement of fracture walls is known to be either predominantly normal or predominantly parallel to the fracture surface [12,21]. Although oblique (opening) displacements are sometimes observed in crack arrays [36], they represent a variation on opening displacement rather than a fundamentally different process [37]. In the simplest sense, perpendicular displacement of fracture walls produces dilatant, mode-I cracks (for extensional normal strains) and anticracks (stylolites; not shown in Fig. 1) for contraction normal to the fracture surface [38]. The former is more commonly encountered in rock engineering than the latter.

Fractures formed during shear deformation — that initially accommodate fracture-parallel displacements — are quite rare in nature and experiments. Many types of fractures (e.g., dilatant cracks, faulted joints) have been misidentified in the literature as such “shear fractures” [12], although sometimes the initial fracture does indeed nucleate with a shearing sense of displacement. Deformation bands [39–41] that form in porous sandstones [42] are one representative of this class of initial fracture. Deformation bands accommodate a continuous shear strain across the narrow, localized band [43], whereas a slip surface that nucleates on de-
formation bands, joints, anticracks, or other surfaces, indicates discontinuous shearing displacements across the surface. However, the terms "shear fracture" or "shear crack" that are entrenched in the literature should perhaps be replaced by more mechanically informative terms such as those noted above (see also Pollard and Aydin [12] for an excellent synopsis of mechanically sound terminologies for cracks). Identifying the superposition of shearing (or other) displacements ontto previously formed fractures is the key to correctly relating fracture type to stress state.

Once one or a number of fractures exists, the rock mass is no longer homogeneous and isotropic in its material properties and the fracture surfaces can be either stronger (Eshelby inclusions) or weaker than their surroundings [15]. Thus dilatant cracks (joints) can yield in shear and slide frictionally (under compressive resolved normal stress) to nucleate faulted joints [32,44–46] (commonly referred to simply as "faults" in the classical geologic literature). On the Mohr diagram this case may be predicted for the point of tangency between the compressional stress state and the shear failure criterion (Griffith, Modified Griffith, Hoek-Brown). Analogously, Willemse et al. [47] demonstrated that anticracks formed in limestone can also localize shear deformation and become faults (faulted anticracks) although the details of fault-zone development in this case differ from that in other rock types. Faults (either deformation bands or slip surfaces) can fail in an extensional sense, producing a dilatant crack superimposed on the previously formed fault. Similar to faulted joints, these jointed faults [31] indicate a change in the stress state, or material behavior, during the deformation and may be used to reconstruct the deformation sequence of the rock mass. All the foregoing examples indicate that an understanding of the displacement sequence and magnitudes accommodated by the fractures themselves is essential to a correct interpretation of rock mass deformation.

2.2. From initial fracture to larger strain examples

Joints are predicted, from a traditional Mohr diagram, when the Mohr circle intersects the failure criterion at the normal-stress axis, on which shear stress equals zero. Pure opening displacements are identified across the fracture (Fig. 2(a), point A), noting that the homogeneous elastic stress described by the Mohr circle is not sufficient as a nucleation criterion for the crack [12,13,20]. If the far-field principal stresses cause the intersection to occur in the tensile regime, but with non-zero shear stresses, oblique dilation of a crack will occur initially [36], producing a mixed-mode crack (Fig. 2(b), point B; [20]), with the sense of shearing mode, I–II or I–III, depending on the 3D state of stress at the fracture tip [48]. Finally, an intersection having a compressive resolved normal stress, and non-zero shear stress by construction, implies frictional sliding along a surface (Fig. 2(c), point C). Technically this stress state and yield criterion predicts faulted joints only, and not the formation of faults having secondary fractures [45], echelon geometries [12,33], and gouge [49,50]. These other attributes are associated with either inhomogeneous stress states in the fractured rock mass or specific processes related to fault slip, neither of which is represented on the Mohr diagram.

Mixed-mode fractures [48,51] have been variously referred to as "transitional tensile joints" (p. 154 [52]) or "hybrid extension/shear fractures" (p. 44 [25]) and [36,53,54]. According to a simple Mohr diagram prediction, such mixed-mode cracks, in an outcrop, might...
be oriented oblique to the remote principal paleostress directions because the cracks would not define the local principal planes. However, this notion is relevant only to the initial fracture and not to its subsequent development into a geologically relevant fracture array. With continued extensional strain, mixed-mode fractures can either produce a fault zone oriented obliquely to the remote stress direction (for compressive normal stresses across the incipient fault zone) or an array of cracks normal to the remote tensile stress trajectory [15]. Horii and Nemat-Nasser [55] demonstrated that the sign of the least principal remote stress strongly influences the subsequent evolution of an initially mixed-mode set of cracks. They showed that a tensile least principal stress promoted continued crack growth as wing cracks [55–58]. These oblique, secondary fractures formed in response to the oblique dilational displacements along the parent mixed-mode cracks and can serve to link individual cracks into larger arrays. Horii and Nemat-Nasser showed that the resulting fracture array would consist of a set of generally parallel dilatant cracks oriented, on average, normal to the tensile remote principal stress.

Mixed-mode fractures are not faults [59], because they do not slide frictionally but open obliquely [36]. The sense of “shearing” mode is understood to be used to describe only the relative lateral displacement, regardless of whether the fracture surfaces are in contact or are separated. It should be understood that shear stresses do not necessarily require shearing displacements: “shearing” displacements, however, necessarily imply frictional contact between the fracture walls. Thus the initial prediction from a Mohr circle, that resolved tensile and nonzero shear stresses imply either fault formation or a mixed-mode joint set is incorrect in the former case, and valid only for the initial displacements and not for a joint set that propagated under this remote stress state to geologically significant strains, in the latter case [26].

The interpretation of fractures associated with the stresses predicted on a Mohr diagram has proven to be a challenge to geologists who may be more familiar with outcrop-scale fracture sets and arrays associated implicitly with much larger strains. However, as pointed out by Jaeger [4] and Hoek [6], among many others, the Griffith criterion, so widely used for tensile-crack predictions, is fundamentally a mathematically convenient simplification of only one of several key processes, including crack nucleation, dilation (where the Mohr analysis applies), propagation, wing-crack growth [56,57], linkage, and arrest or termination [12,55,58]. Because the Mohr-circle approach — using either Coulomb, Griffith, Modified Griffith, or Hoek-Brown criteria — is useful only for resolving stresses onto a particular plane of given orientation, it is perhaps best suited for understanding the initial phase of crack opening (but not its nucleation). Further mechanical development of a fracture set, array, or network in rock, leading to geologically interesting examples at larger strains (pp. 25–28 [7,49]) and also [26,35], requires more complete treatments of the processes noted above (pp. 42–43 [25]) and [47,55,57,60–62].


Natural joint patterns are characterized by systematic variations in fracture length and spacing that relate to the conditions under which the fractures grew. In this section joint sets are considered that developed largely free of the confines of stratigraphy (e.g., bedded sedimentary rocks) that can lead to evenly spaced joints [21,25,63]. Mode-I fracture sets developed in granitic plutons or in massive sedimentary rock (without the stratigraphic control in either case) demonstrate the dependence of fracture population statistics on the smaller scale processes at the fracture tips.

Several crack growth simulations, performed numerically, have adopted the Charles [64] law as an approximate model for subcritical crack growth in rock [60,61,65,66,66]. The mechanism of stress corrosion [67,68] used in this approach assumes that fracture propagation is facilitated by chemically assisted weakening of atomic bonds in the near-tip region [68]. Stress corrosion is considered to be the dominant mechanism for subcritical crack growth in the Earth’s crust [65], with other mechanisms such as diffusion, dissolution, ion exchange, and microplasticity requiring special conditions of high temperature or reaction rate [68].

The Charles law is given by [68]

$$v = v_0 \exp\left(-\frac{H}{RT}\right)K^n$$

in which $v$ is the velocity of fracture propagation (usually for mode-I), $H$ is the activation enthalpy [69], $R$ is the universal gas constant, $T$ is absolute temperature, $K$ is the stress intensity factor, $v_0$ is the threshold velocity for subcritical crack growth, and $n$ is the stress corrosion index. Conditions appropriate for subcritical crack growth of tectonic fracture arrays are found associated with chemically active, corrosive environments perhaps typical of the Earth’s crust (e.g., sedimentary basins and petroleum fields, hydrothermal systems, ore deposits, active fault zones). Typical values for $n$ in rock [70] are approximately $15 < n < 100$, with lithology, anisotropy, temperature, confining pressure, and chemical environment being major controls on this exponent. The value of $n$ also increases with the complexity of rock-mass structure, indicating a greater re-
sistance of poly lithologic rock types to subcritical crack growth relative to monomineralic rocks under the same environmental conditions [68].

Subcritical crack growth depends on the local stress state, rock type, the relative chemistries of fluid (usually water) and rock mass, and a reaction rate at the fracture tip that exceeds the overall strain rate. At sufficiently low rates of tectonic strain, the rate of propagation of the fracture tip through the rock mass is limited (and governed) by the chemical reaction rates, leading to subcritical crack growth (Fig. 3, region I). As the tectonic strain rate exceeds the reaction rate-limited values, however (Fig. 3, region II), the chemical processes at the fracture tip cannot keep pace with the applied strain, leading to the more familiar quasi-static crack growth governed by fracture toughness ($K_r K_c$ [13,51]; Fig. 3, region III). The remote-load boundary conditions (fixed-grip or dead-weight loading) determine whether the LEFM crack propagates stably and quasi-statically or accelerates into a dynamic “running” crack [11,51,66,71].

The subcritical crack growth regime is characterized by positive values of $n$ and a proportional relationship between crack velocity and the stress intensity factor. In principle, as $n$ approaches 0, the slope becomes horizontal in this region on Fig. 3 and subcritical crack growth is not an appropriate mechanism. A horizontal slope implies that crack velocity becomes independent of $K$, as in quasistatic crack growth under LEFM conditions. Values of $n \leq 1$ are sometimes taken to represent an approximation to quasistatic behavior in numerical modeling studies [60] (J. Olson, pers. comm., 1998).

The Charles law can be recast into the simplified form [60,65]

$$v = v_0 \left( \frac{K}{K_{\text{max}}} \right)^n$$

in which $K_{\text{max}}$ is defined by the longest crack in the array. This expression may also be written as the

$$v = v_0 \left( \frac{K}{K_{\text{max}}} \right)^n$$

Fig. 3. General relationship between stress intensity factor and propagation velocity of the fracture tip [after 67,68]. Mechanisms controlling crack growth in the subcritical regime ($K_0 \leq K \leq K_c$) are listed (I, II, III). For $K \geq K_c$, LEFM implies either quasistatic propagation (for fixed-grips, displacement-controlled loading) or dynamic propagation (for dead-weight, stress-controlled loading).

Fig. 4. Histograms of fracture length for four crack-growth simulations, after Olson [60]. Upper panels, quasistatic growth ($n = 1$); lower panels, subcritical crack growth. Note variable vertical scales (left column) and log scales (right column, same data).
equivalent, energy-based form obtained by using the $K$-based expressions for fracture propagation energy $G$ \cite{10,11,61}. Olson \cite{60} and others have used this form of the Charles law to numerically model the growth of mode-I dilatant crack arrays. By varying the value of $n$ and allowing the crack array to propagate and develop under displacement-controlled conditions, a variety of fracture network geometries was achieved. This work shows \cite{60,61} that the length–frequency, and spacing, of fractures in an array are sensitive to the propagation criterion at the fracture tip in general, and to the degree of subcritical crack growth (through the stress corrosion index $n$) in particular.

As an example, the length–frequency data from synthetic fracture arrays generated by Olson \cite{60}, displayed as histograms in Fig. 4, demonstrate that the stress corrosion index $n$ exerts a significant influence on the sorting of fracture lengths (and spacings \cite{60}) as a fracture network grows through a rock mass. As $n$ is increased, the number of longer fractures is reduced, leaving a greater number of short, initial microcracks that have not propagated significantly far. Joint clustering is also observed between the few long fractures \cite{60,61} at larger values of $n$, leading to specific channels of high hydraulic conductivity though an otherwise impermeable rock mass \cite{10}. Even though the value of normal strain accommodated by each simulated array is comparable, the degree of mechanical interaction and linkage changes as subcritical crack growth becomes important. Many natural joint sets are similar geometrically to the numerical simulations reported by Olson \cite{60} and Renshaw and Pollard \cite{61} such as those in Sierran granites documented by Segall and Pollard \cite{29}.

The histograms of fracture length, and related observations of fracture spacing \cite{60}, also suggest that subcritical crack growth influences the strength and deformability of the rock mass. Fracture spacing correlates directly with Rock Mass Rating (RMR) (or the equivalent Q-system value, through the calculations of equivalent RQD and fracture frequency) \cite{5}, and, mechanically, with the deformation modulus and shear strength of the rock mass \cite{6,7,11,15,18}. The numerical simulations show \cite{60} that fracture sets growing under conditions more analogous to LEFM ($n \leq 1$) may be characterized by a relatively close spacing, whereas fracture sets that grew subcritically ($n > 1$) have many short (less persistent) fractures between a few long fractures that are separated by wide spacings. In prac-

![Fig. 5. Cartoon showing differences in fracture network geometry [after 60,61]. (a) Quasistatic growth; (b) and (c), greater values of stress corrosion index. Vertical bars represent possible scan-line traverses designed to assess fracture spacing for RMR.](image)

![Fig. 6. Plot of cumulative frequency vs length for synthetic fracture populations [60]. Shaded regions indicate portions of curves where the statistics are unreliable due to low numbers of fractures (right-hand side) or simulation artifacts (fixed number of initial flaws = 60, left-hand side).](image)
tice, a scan-line traverse [3,8] through a rock mass having $n > 1$ may miss recognizing and counting the shorter fractures, leading to elevated values of equivalent-RQD and RMR [5], and therefore larger values of rock mass strength (unconfined compressive, tensile, and shear) and deformability.

The domain size may also increase for a “subcritical” fracture set (Fig. 5) if it is defined as the distance between the largest fractures that cross the traverse measuring tape. Alternatively, the domain size could become even greater if the engineer recognizes that the largest fractures are merely members of a still larger-scale fracture network. Thus the definition of domain size, required for relating outcrop characteristics to those of an effective continuum for modeling purposes, depends on a realistic assessment of fracture spacing (or density along a scan-line) and on the recognition by the rock engineer of fracture population geometries. Improved methods for calculating fracture density [9] may provide more robust and informative inputs into rock mass classification systems than those presently available using either scan–line or box–counting techniques.

The histograms of fracture length–frequency [60] are compared in Fig. 6 using the standard [72,73] cumulative-frequency diagram for fracture population analysis. The curve for the baseline fracture set produces a continuously curving distribution, suggesting that a linear fit to the data, and therefore a power-law distribution of fracture lengths, is not appropriate for this numerical simulation. However, as $n > 1$ the overall slopes of the distributions appear to decrease and become more nearly linear, consistent with the power-law distribution of fracture lengths (from $2 < L < 6–7$ m, where the statistics are more robust; unpatterned area in Fig. 6) noted by Olson [60] for his subcritical growth simulations and by Segall and Pollard [29] for joint sets in Sierran granite. While specific relationships between the stress corrosion index $n$ and the slope of the cumulative-frequency population exponent cannot be fully determined from these few numerical experiments, a physical relationship between the two appears reasonable given the clear dependence of incremental fracture length–frequency [73] and $n$ demonstrated above (Fig. 4). The results are sufficiently encouraging to motivate coupled investigations of near-tip processes, as a test of subcritical crack growth, and the fracture population statistics.

4. Conclusions

Consideration of discontinuous fracture geometries leads to a more coherent nomenclature for cracks and faults in rock engineering and sounder relationships between common peak-strength failure criteria and their associated fracture arrays developed at larger geologic strains. Many problems in fracture terminology can be avoided by recalling their specific strain implications. The growth and development of natural tectonic fracture sets in rock are modulated by the mechanical and chemical processes operating at the fracture tip. These processes serve to couple the focused inhomogeneous stress states generated around fracture tips to the more continuously deforming rock mass through mechanisms that are sensitive to stress state, strain rate, and environmental conditions (temperature, relative chemistry of rock and fluid phases). Because the characteristics of the fracture tips strongly influence the nature of fracture propagation, interaction, and linkage, the overall geometric, mechanical, and hydrologic properties of the ultimate fracture population must also depend on the processes acting at the fracture tips.

Acknowledgements

I thank John Kemeny, John Hudson, and Jon Olson for helpful informal comments on aspects of the work presented here, and Larry Myer and Chin-Fu Tsang for organizing the enjoyable Cook Conference at LBL where the paper was refined. Thoughtful review comments improved the final paper. This work was supported in part by grants from NASA’s Planetary Geology and Geophysics program and completed while the author was enjoying a pleasant sabbatical stay at Woods Hole Oceanographic Institution and at the University of Paris. I thank Jian Lin for his hospitality and intellectual collaboration at WHOI and Daniel Mége for the same in Paris.

References


