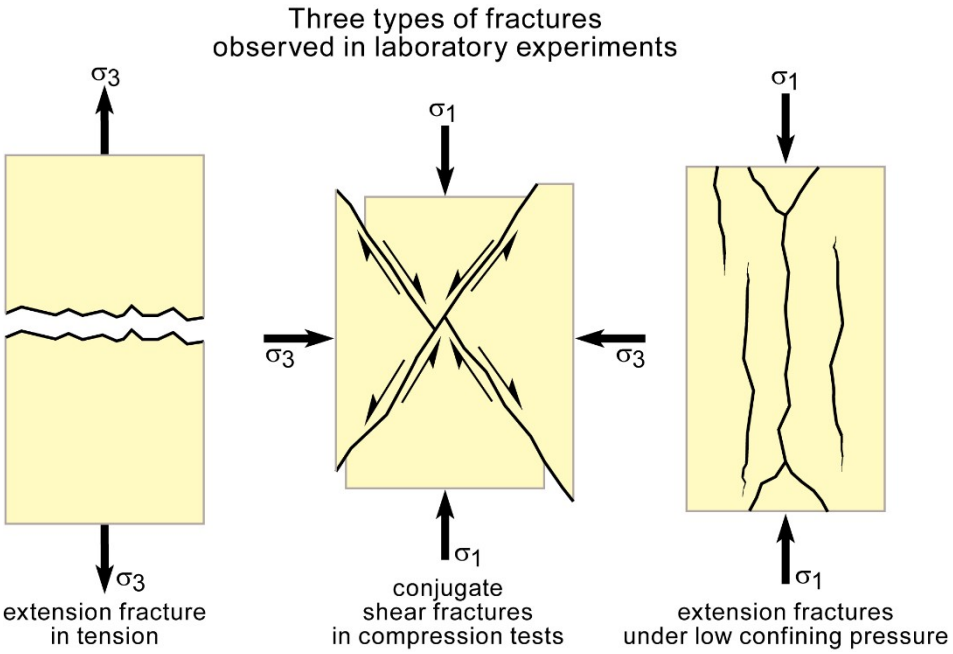
BRITTLE FAULTING

Because most rocks are brittle at low temperature and low confining (lithostatic) pressure, virtually every rock at or near the Earth's surface exhibits evidence of brittle failure, i.e. deformation-induced loss of cohesion.
Brittle failure results from the irreversible and very rapid propagation and connection of **cracks**, a process called **fracturing**. Cracks are grain-scale planes that pre-exist or nucleate under stress on some microscopic defect of the material. Fracturing initiates at stress levels near the **yield strength** (or maximal differential stress σ_1-σ_3), marking the elastic limit of the material. Tensile stress cannot act across the resulting physical discontinuity that has no cohesion.
By reference to laboratory observations, a broad mechanical classification of fractures recognizes two fracture modes:

- **Extension fractures**, resulting from the initial separation of two formerly contiguous surfaces; displacement is parallel to the minimum principal stress (i.e. maximum tension), which is orthogonal to the fracture plane.
- **Shear fractures** resulting from initial displacement along the fracture plane; fractures and displacement are oblique to the maximum principal stress (maximum compression).

This lecture discusses qualitatively the relationships of the earthquake source to faulting.



Brittle faulting is a process producing localized offset along a shear fracture. This lecture discusses how and qualitatively comments the relationships of the earthquake source to brittle faulting.

Fault orientation relative to principal stress axes

Theoretical considerations

A force F that acts on P can be resolved into components normal (F_N) and parallel (F_S) to the plane P. The components have magnitudes:

$$F_N = F\cos\theta \qquad \text{and} \qquad F_S = F\sin\theta \qquad (1)$$

with θ the angle between the direction of the force and the normal to the plane (see lecture on mechanical aspects of deformation; forces and stresses).

Drawing cross-sections of a cube with the force of magnitude F acting normally to one cube face of area A, the stress is by definition the concentration of force per unit area, which can be visualized as the intensity of force if the cube is such that its faces = unit area = 1. Thus, the magnitudes of the normal and shear components of stress across a plane P are:

$$\sigma_N = F_N / A_P = (F/A)\cos^2\theta = \sigma\cos^2\theta$$

and (2)

$$\sigma_S = F_S / A_P = (F/A)\sin\theta\cos\theta = \frac{\sigma}{2}\sin 2\theta$$

Typically, any rock is under a triaxial state of stress, and σ_1, σ_2 and σ_3 are the principal stresses with $\sigma_1 \geq \sigma_2 \geq \sigma_3$.

Remember! The convention in geology takes all positive stresses as compressive. In the non-geological literature, extensional stresses are positive!

For practical purposes, one considers an arbitrary plane P within the body, parallel to σ_2 and whose normal line makes an angle θ with σ_1 (θ is also the angle between the plane P and σ_3). One assumes that, for an elementary treatment of the state of stress and of the quantitative relationship between the normal and shear stresses, one may neglect σ_2 and only consider the two-dimensional principal plane (σ_1, σ_3). Then, lines in that plane represent traces of planes perpendicular to it and parallel to σ_2.
All considered planes will make a line in this plane.
The stress tensor can be represented by its two principal components σ_1, and σ_3. Where the principal stresses are σ_1 and σ_3 the equations for the normal and shear stresses across a plane whose normal is inclined at θ to σ_1 are:

$$\sigma_N = \frac{(\sigma_1 + \sigma_3)}{2} + \frac{(\sigma_1 - \sigma_3)\cos 2\theta}{2}$$
$$\sigma_S = \frac{(\sigma_1 - \sigma_3)\sin 2\theta}{2} \qquad (3)$$

The principal stresses σ_1 and σ_3 can be assumed to represent the tectonic stress field (for instance, compression σ_1 is horizontal and σ_3 vertical). With this assumption, the equations (3) are important because they can be applied to relate regional tectonic stresses to the normal and shear stresses on local fault planes.
Equations (3) demonstrate that the value of σ_S is maximum when $\sin 2\theta = 1$ i.e. $2\theta = 90°$. Thus, the planes of **maximum shear stress** make a theoretical angle of 45° with σ_1 and σ_3. The maximum shear stress has the value $(\sigma_1 - \sigma_3)/2$.
In all cases where $\sigma_1 \geq \sigma_2 \geq \sigma_3$ there are two planes of maximum shear stress. The paired faults, called **conjugate faults,** develop more or less synchronously in both of the equally favored orientations; conjugate faults intersect in a line parallel to the intermediate principal stress axis σ_2.
In the special case where $\sigma_2 = \sigma_3$ or $\sigma_1 = \sigma_2$, there is an infinite number of planes inclined at 45° to σ_1 or σ_3. All possible orientations are tangents to a cone.
Equations (1) and (3) imply that compressive normal stresses tend to inhibit sliding while shear stresses tend to promote sliding on any plane. This lecture will now discuss how these conclusions apply to brittle deformation, keeping in mind that the demonstration is valid for the onset of faulting only.

Shear fracture criteria

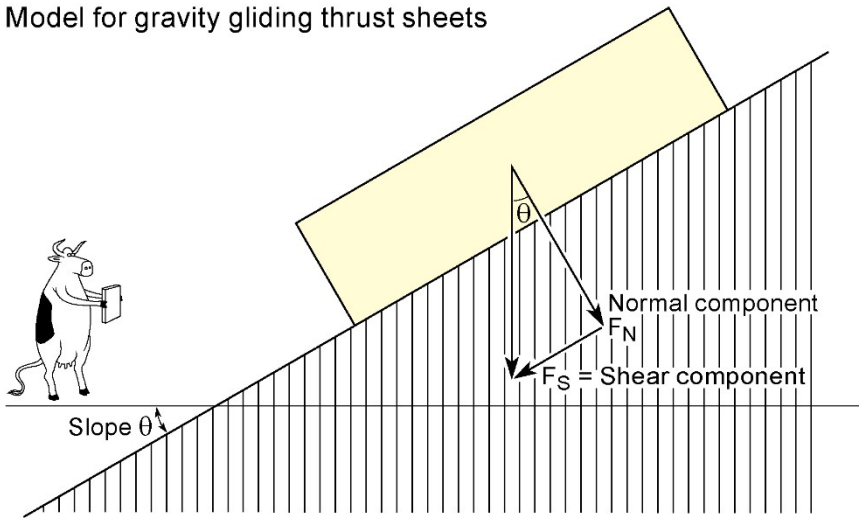
Shear failure occurs when loading creates shear stresses that exceed the shear strength of the rock. A yield criterion is a hypothesis concerning the limit of elasticity under any combination of stresses. Three main failure criteria are applied to rocks:

- The Coulomb criterion;
- The Mohr envelope;
- The Griffith crack theory;
- None of the basic "friction laws" discussed in the following paragraphs is based on mechanical processes. They are only phenomenological, in-equation descriptions of experimental faulting. This limitation is due to the complex and constant evolution of surface contacts during fault displacement (roughness changes with asperities being sheared, plowing, stepping up and down, the interlocking of asperities, and strength of the rock to be sheared and chemical reactions producing **wears**).

Frictional sliding

Three coefficients refer to friction: (1) **internal friction**, to create a sliding surface;
(2) **static friction** to initiate movement on the sliding surface;
(3) **dynamic friction** to maintain sliding on the surface.

First investigations considered simple experiments with objects sliding down an inclined plane under the action of gravity.



Only two forces were considered: the vertical weight of the object W and the horizontal force F, which at rest are balanced by opposite reactions W_r and F_r, respectively. F_r arises from frictional forces that resist the motion that F would impel. At some magnitude F_r breaks and the object moves. According to the weights of different objects, some proportionality between the two forces was established:

$$\text{If } F = F_r < \mu_s W \text{ the object does not move, it is static.}$$

$$\text{If } F = F_r > \mu_s W \text{ sliding initiates, the object is accelerated.}$$

This is the Amonton law, in which the **static friction** μ_s is independent of the area of contact.

In reality, the shear component of weight W_S acting parallel to the inclined plane initiates sliding while the normal component W_N is the resisting force. The static friction μ_s becomes the ratio of the shear stress to the normal stress:

$$\mu_s = \frac{W_S}{W_N} = \frac{\sin\theta}{\cos\theta} = \tan\theta$$

where θ is the inclination of the plane termed **angle of friction**.
Such observations lead to the concept that two bodies with a plane surface of contact are pressed together by the normal stress σ_N. The shear stress σ_S necessary to initiate sliding is related to σ_N, by an equation of the form:

$$\sigma_S = f(\sigma_N) \tag{4}$$

which, reformulating the Amonton law, becomes:

$$\sigma_S = \sigma_N \tan\phi = \mu_s \sigma_N \tag{5}$$

ϕ is the **angle of internal friction**. It is the angle between the vector normal to the rupture plane and the stress acting on this plane. The term internal friction is not strictly friction. It describes a material property of slip resistance along the fracture and the state of the surfaces in contact (e.g. smooth versus rough, fresh versus altered or coated, etc.). It also involves the coalescence of microcracks. Experiments have shown that this linear relationship is valid for materials with no cohesion strength, such as soils.

Coulomb criterion

Charles Augustin de Coulomb found that there is both a stress-dependent and stress-independent component of shear strength. He proposed in 1776 that shear fracture occurs when the shear stress along a potential fault plane overcomes two forces: (1) the cohesive strength of the material before failure on that plane and (2) the resistance along that plane once it had formed. He expressed shear resistance S as:

$$S = c.a + \frac{1}{n} N$$

where c is the cohesion per unit area, a is the area of the shear plane, N is the normal force on the shear plane and $1/n$ **is the coefficient of internal friction**.
In modern terms, this equation is written:

$$\sigma_S = c + \mu \sigma_N \tag{6}$$

where c is the **cohesion**, a material constant;
μ is the coefficient of internal friction, another material constant equivalent to $\tan\phi$, the coefficient of internal friction seen for cohesionless soils (equation 5).

Both parameters are not inherent properties of the tested material but depend on the test conditions. Equation (6) is often referred to as the **Coulomb criterion**.

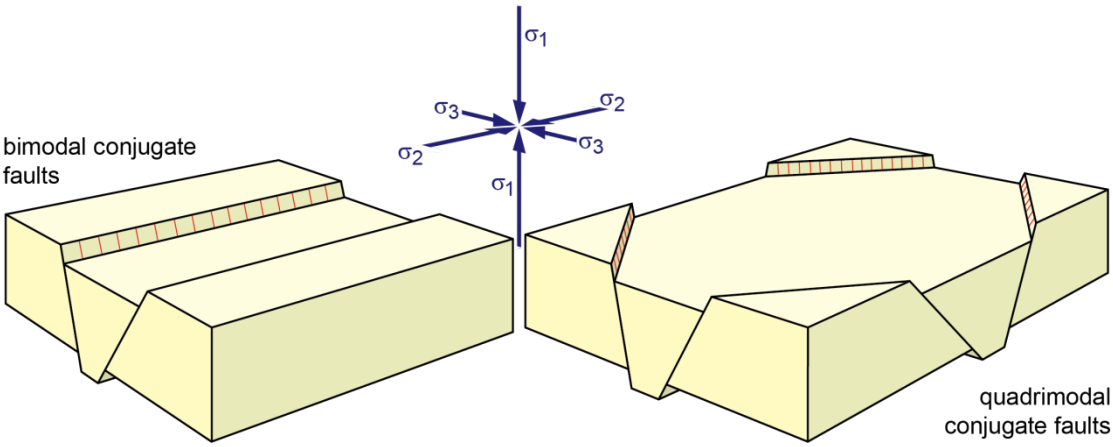
Physical interpretation

Equation (6) assumes that shear fracture in solids involves two factors together:
- Breaking cohesive bonds between particles of intact rock (the c term); cohesion is a measure of this internal bonding.
- Frictional sliding (the term μ, proportional to the normal compressive stress σ_N acting across the potential fracture plane); internal friction is caused by contact between particles.

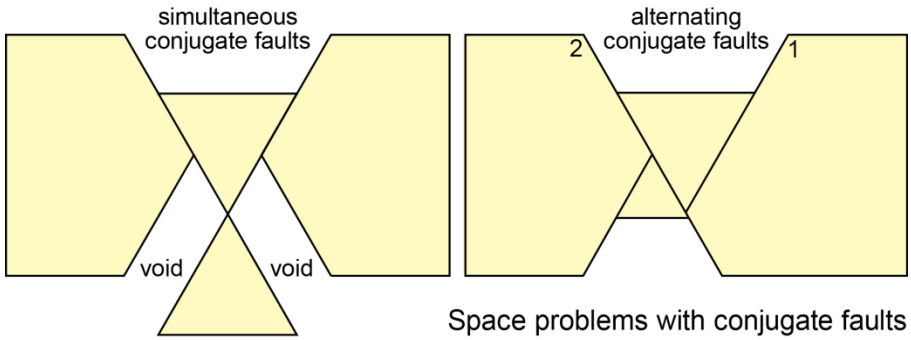
This physical interpretation provides an acceptably good fit with many experimental data, which yield cohesive strength of 10-20 MPa for most sedimentary rocks and about 50 MPa for crystalline rocks. The average angle of internal friction is 30° for all.

Experimental fractures

In triaxial experiments (σ_1, σ_2 and σ_3 have non zero magnitudes) the actual shear fractures form **dihedral angles** smaller than 45° to the maximum compressive stress axis σ_1. A good average for rocks is approximately 30°. The acute bisector of conjugate faults is parallel to σ_1. This geometrical relationship between the stress axes and the shear fractures is widely utilized to define in the field the orientation of the stress axes from measurements of faults. Understanding faulting implies understanding the difference between this ideal angle and the fault direction. The factors that contribute to this angular difference are included in the concept of "**angle of internal friction**".
Three-dimensional deformation involves the activation of polymodal sets of fractures, usually comprising variously oriented sets of conjugate (bimodal) faults. For instance, a pair of conjugate sets forms a quadrimodal fault set. These quadrimodal faults intersect to form rhombohedral traces on outcrop surfaces with σ_1 and σ_2 bisecting the acute angles between the fault planes.



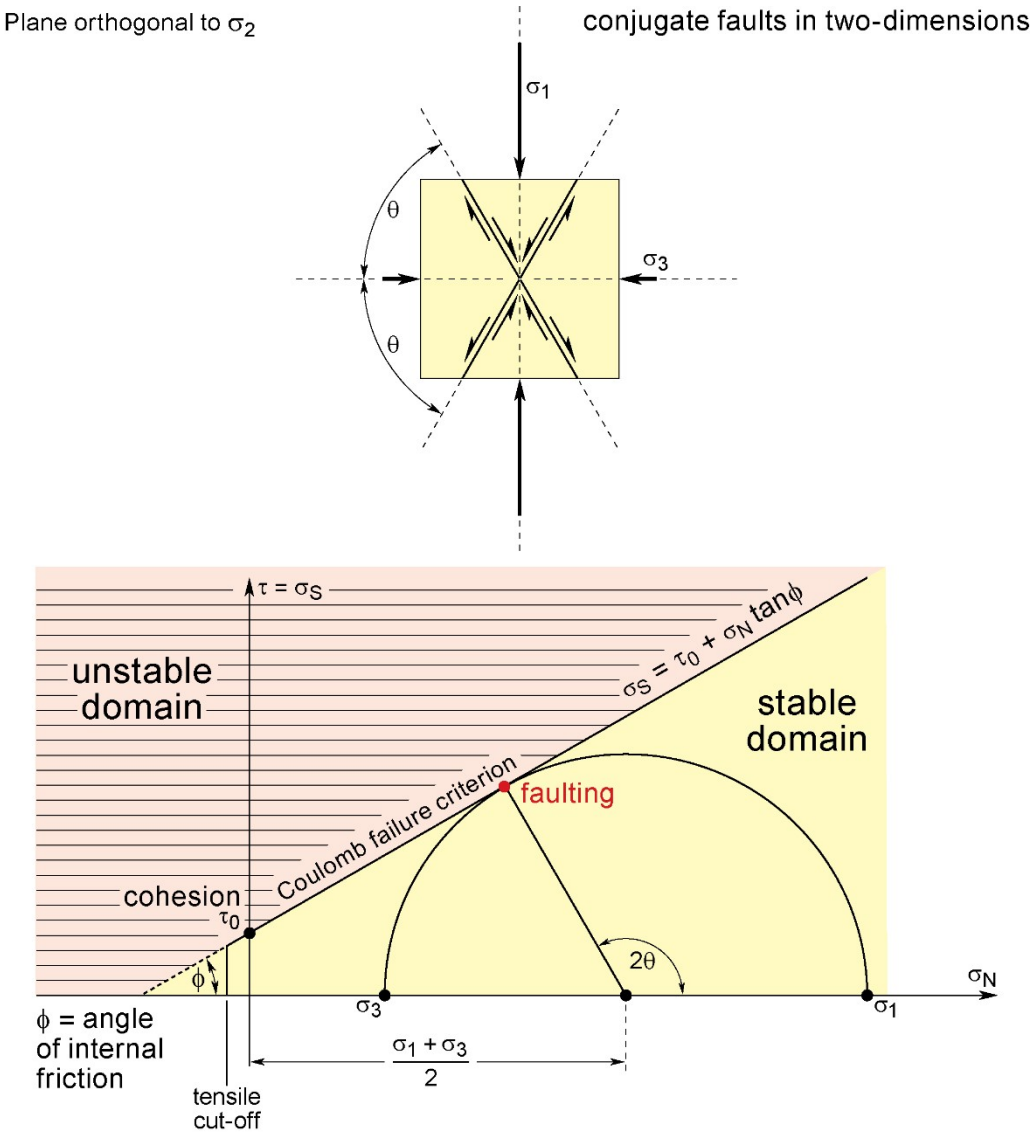
The notion of simultaneous conjugate faults is geologically valid, provided "simultaneous" loosely means alternating over a short amount of time. Space problems ensue from strict simultaneity; they can be solved only by rotation and alternating differential slip on each of the conjugate faults.



Graphical representation

The general form of equation (6) also predicts that failure points in a σ_S/σ_N diagram should lie on a straight line with slope μ and intersecting the σ_S ordinate at the shear strength $c = \tau_0$. The line is characteristic of many rocks tested at moderate confining pressures and describes their shear failure to a good approximation.
The graphical representation shows that:
- Any stress condition below the Coulomb criterion line is safe, but shear failure occurs as soon as the stress condition touches the line. Then slip relax stress, thereby pre-empting stress conditions that would exceed the failure line (i.e. Mohr circles growing beyond the criterion line).

- The shear fractures form at less than 45° to σ_1 since the shearing resistance line has a positive slope. This slope represents the competing effects of σ_N and σ_S on shear fractures. Both the minimum normal stress and maximum shear stress promote together new shear fractures. The acute angle between conjugate faults in rocks (60 to 70°) is an optimization of these two conditions.



Two-dimensional representation of the Coulomb failure criterion with Mohr stress circle at failure

- Since slip depends on the magnitude and not the sign of shear stress, there are two lines symmetrical with respect to the horizontal axis of normal stresses. These lines have slope angles $\pm\phi$. Accordingly, the acute angle between the conjugate faults is bisected by the greatest principal stress σ_1.
- The criterion is linear but since rocks cannot sustain large tensile stresses a tension cut-off (a vertical line on the negative side of normal stresses) is often introduced.

The graphic shows that the angle of internal friction ϕ, which is a material property, is related to the angle θ of the fault plane: $\theta = \frac{\phi}{2} + 45$. The linear failure criterion imposes that new shear fractures make with σ_1 a well-defined angle of $\pm(45° - \phi/2)$.

Linear relationship between σ_1 and σ_3

The graphical representation of a Mohr circle shows that the radius is:

$$\frac{\sigma_1-\sigma_3}{2}=\left(\frac{\sigma_1+\sigma_3}{2}+\frac{c}{\tan\phi}\right)\sin\phi$$

Multiplying both sides by 2 and transposing terms, this becomes:

$$\sigma_1\left(1-\sin\phi\right)=\sigma_3\left(1+\sin\phi\right)+2c.\cos\phi$$

Expressing σ_1:

$$\sigma_1=\frac{2c.\cos\phi}{1-\sin\phi}+\sigma_3\left(\frac{1+\sin\phi}{1-\sin\phi}\right) \qquad (7)$$

Playing with the classical sine and cosine identities $\sin\alpha=\pm\sqrt{1-\sin^2\alpha}$ and $\cos\alpha=\pm\sqrt{1-\sin^2\alpha}$:

$$\frac{\cos\phi}{1-\sin\phi}=\sqrt{\frac{1+\sin^2\phi}{(1-\sin\phi)^2}}=\sqrt{\frac{(1-\sin\phi)(1+\sin\phi)}{(1-\sin\phi)(1-\sin\phi)}}=\sqrt{\frac{1+\sin\phi}{1-\sin\phi}}$$

Substituting this result into equation (7) gives:

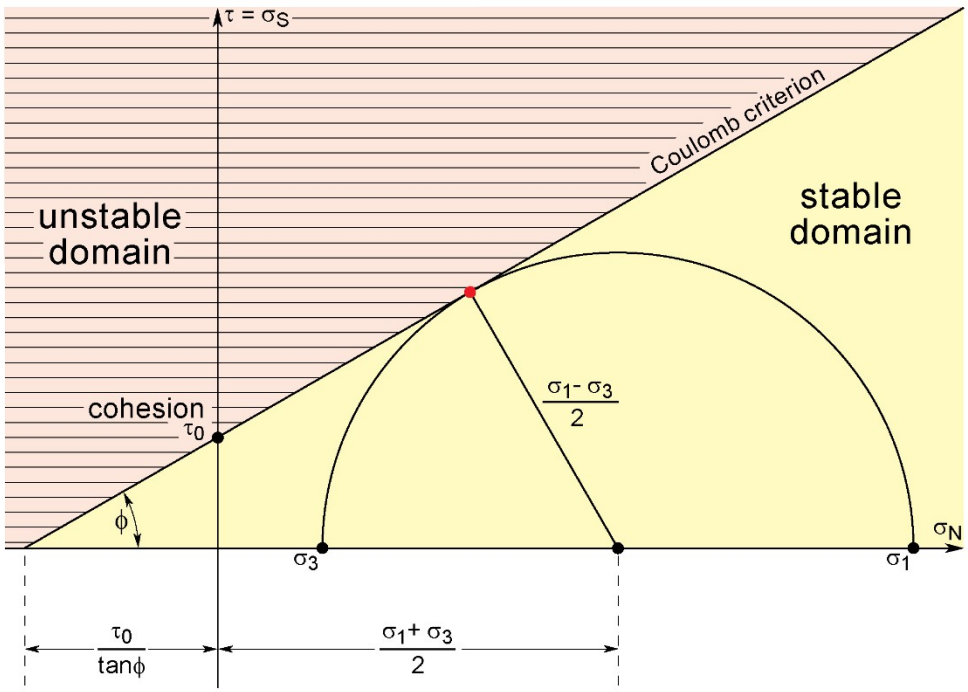
$$\sigma_1=2c\sqrt{\frac{1+\sin\phi}{1-\sin\phi}}+\sigma_3\left(\frac{1+\sin\phi}{1-\sin\phi}\right)$$

Since c and ϕ are material constants, this alternate expression of the Coulomb criterion is of the linear form:

$$\sigma_1=a+b\sigma_3$$

$$a=2c\sqrt{b} \qquad \text{und} \qquad b=\frac{1+\sin\phi}{1-\sin\phi}$$

This relation shows that σ_1 and σ_3 are linearly related when a fracture occurs.



Relationship between σ_1 and σ_3 at rupture in a two-dimensional Mohr diagram

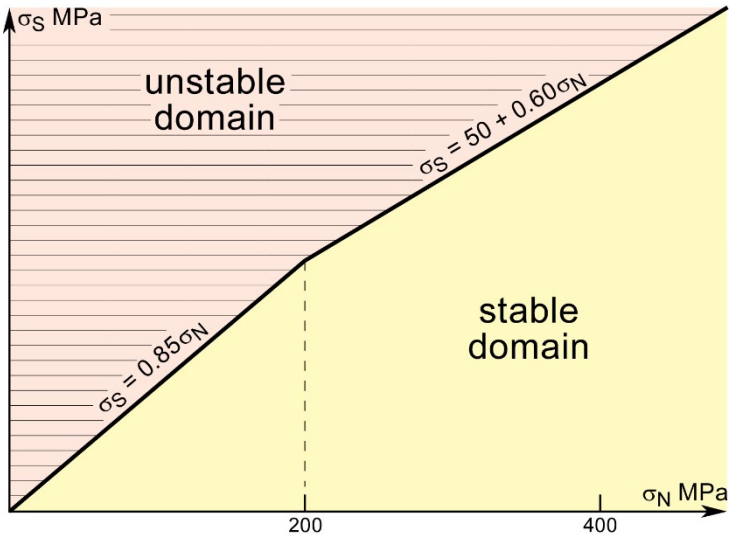
Real-world: Byerlee law

Compilation of experimental data on samples with pre-cut fault surfaces indicates that **frictional sliding** (slip of pre-existing fracture planes) is independent of rock type, except for many clay-rich rocks. Resistance to shearing depends on confining pressure σ_N. Two best-fitting, general and empirical equations expressing this relation are known as Byerlee's law:

For $\sigma_N < 200\text{MPa}$ $\quad \sigma_S = 0.85\,\sigma_N$ (8)

For $200\text{MPa} < \sigma_N < 2000\text{MPa}$ $\quad \sigma_S = 0.6\,\sigma_N + 50\,\text{MPa}$ (9)

These laws were established at room temperature but emphasize that friction is proportional to the normal stress, which presses fault blocks together.



Byerlee's law for pre-existing fault surfaces

Mohr conditions

Christian Otto Mohr proposed in 1900 that σ_S and σ_N are related in general by a non-linear function specific to the material considered. The function represents the shape of the envelope to a series of Mohr circles at failure for a given material: the **Mohr envelope**, which is an empirical curve that delimits the field of failure of the material.

Form of the Mohr envelope; graphical representation

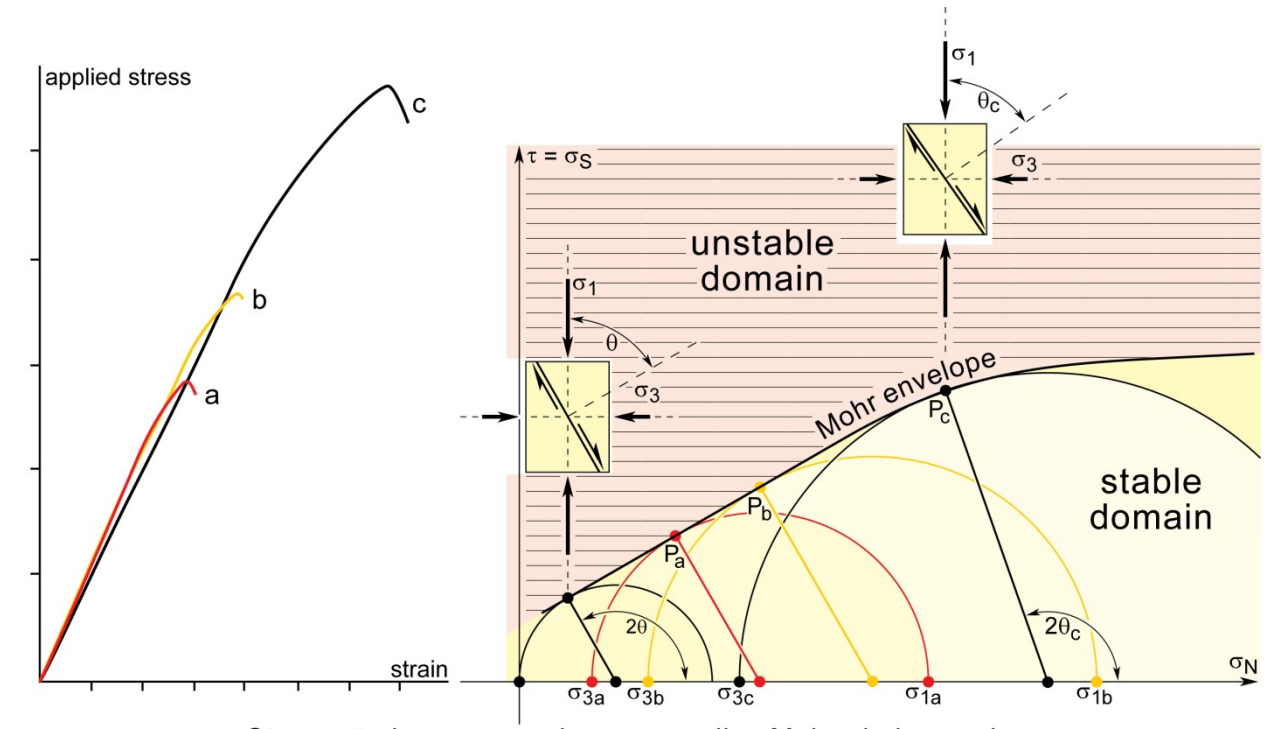
The Mohr envelope is slightly concave towards the σ-axis and is symmetric with respect to this axis. Its shape and its position vary for each material and are empirically obtained as follows:

- A cylinder of rock is axially compressed under constant confining pressure. The axial load is gradually increased until the rock fails. The Mohr circle at failure contains a point P, which represents the σ_S and σ_N stresses on the failure surface and indicates the orientation of the planes along which the rock has failed, at an angle θ to σ_3 (here the confining pressure).
- 2θ is the trigonometric angle between the radius normal to the tangent to the Mohr circle at P and the σ-axis.
- Several experiments are performed on identical samples of one rock type at different confining pressures. They yield slightly different Mohr circles. The tangent to these circles, which passes through all P points, represents the Mohr envelope or **failure envelope** for the particular rock.

A series equation expresses this envelope:

$$\sigma_1 = \frac{2c\cos\phi + \sigma_3(1+\sin\phi)}{1-\sin\phi}$$

Drawing the symmetrical failure criterion below the axis of normal stresses sets conjugate shear fractures at angles $-\theta$ to σ_3.



Stress-strain curves and corresponding Mohr-circle envelope
experimentally defined by faulting (failure) in identical rock samples at various confining pressures

The shape of the Mohr envelope also illustrates that:

- The curve cuts the normal stress axis only at one point, which means that it is impossible to cause shear fractures with hydrostatic pressure.
- The shear stress required to produce failure increases with the confining pressure.
- The Mohr circles at failure become progressively larger with the size of the confining pressure.
- At high confining pressures, the envelope delineates a symmetrical pair of parallel and horizontal lines at a critical **shear** strength (Von Mises criterion), which means that the material becomes perfectly plastic, while the ductile flow is pressure-insensitive.
- The progressively decreasing slope with increasing pressure to horizontal at high pressure represents the progressive transition from brittle to ductile behavior.

At intermediate confining pressures, the fracture strength usually increases linearly with increasing confining pressure. The angle between this line and the horizontal axis is the angle of internal friction ϕ and the slope of the envelope is the Coulomb coefficient μ with, as above

$$\mu = \tan\phi$$

In soil mechanics, the curved envelope is considered as a line.

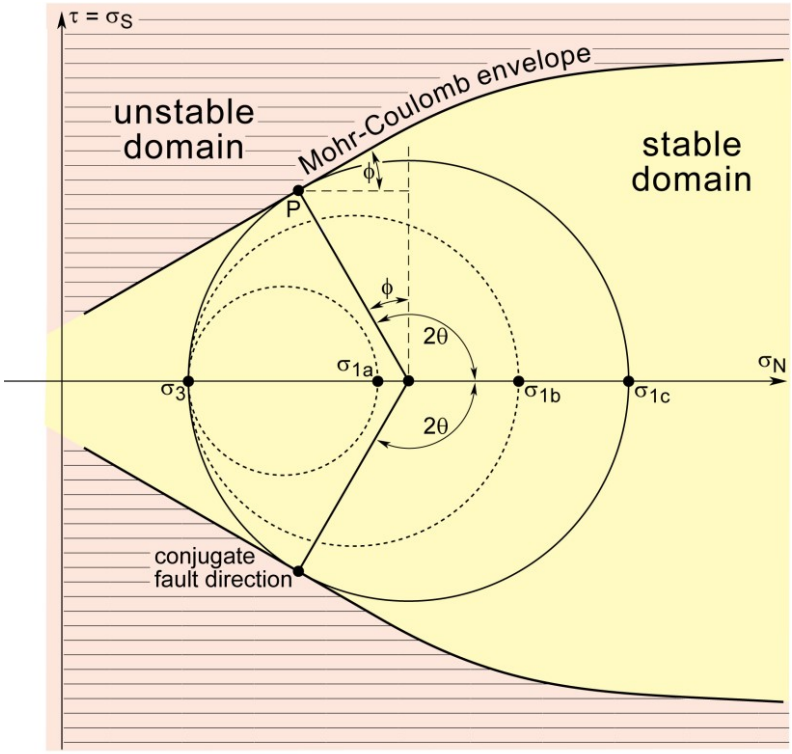
Physical concept and interpretation

The curvature of the Mohr envelope is attributed to the increasing proportion of cracked compared to intact areas on the incipient rupture surface. The strength of this incipient plane is then a combination

of the frictional resistance across the cracked areas and some measure of the bulk cohesion strength of the material in the intact areas.
Once a Mohr envelope is established on a given rock type, this curve can be used to predict both the **ultimate strength** and the fault angle in tests at other confining pressures. The shearing resistance can be calculated with the equation:

$$\sigma_S = \frac{(\sigma_1 - \sigma_3)\sin 2\theta}{2}$$

2θ is the trigonometric angle between the radius through P (normal to the tangent to the circle) and the σ-axis.



Orientation of new conjugate faults
fitting the Mohr-Coulomb criterion in a Mohr diagram
The dashed Mohr circles represent successive states of stress without faulting

A simple geometrical construction shows that the angle 2θ in a Mohr construction must exceed a right angle by a value of ϕ, which expresses the angle of internal friction. The consistent orientation of new shear surfaces corresponding to Mohr-Coulomb failure is:

$$2\theta = \pm(90° + \phi)$$

The sign $\pm$ implies that there are two symmetrically oriented (conjugate) fault directions.

The two-dimensional Mohr-Coulomb failure theory assumes that failure is only a function of the differential stress $(\sigma_1 - \sigma_3)$, i.e. the diameter of the Mohr circle. σ_2 has no influence.

The law is based on the following concept:

- If a Mohr circle representing a particular combination of σ_1 and σ_3 does not intersect the envelope, the material will not fracture and remains elastic.
- If the Mohr circle touches or intersects the Mohr envelope, the material will fracture. The contact point defines the orientation of the fracture plane. Note that tangency of the circles to the Mohr envelope determines the inclination of the failure plane that will form: only one angle θ is possible at failure.

- No part of any Mohr circle can go above the envelope, into the failure field, because the critical stresses are exceeded.

In practice, only a limited portion of the envelope is available; the hydrostatic tension (negative stresses) has not been attained experimentally.

Griffith criterion

Physical background

Alan Arnold Griffith noticed that strength values measured on glass in tension are smaller than those predicted by solid-state theory.

Breaking inter-atomic bonds

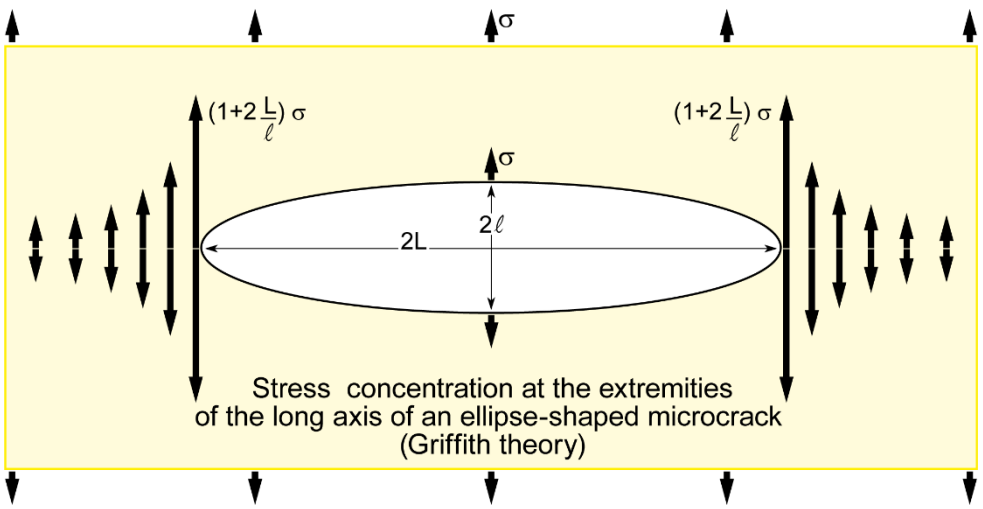
The solid-state theory considered the atomic scale, with a crack developing through breaking at once all inter-atomic bonds across the whole crack surface. Knowing the strength of single bonds, the calculated strength to break all the bonds is 10 to 1000 times bigger than the experimentally measured tensile strength. The discrepancy between theoretical and empirical strength is explained with microscopic flaws or cracks found either on the surface or within the material, and these cracks produce local stress magnification at their tips. The fracturing of material has to start somewhere within the material. It usually starts at a location of stress concentration.

Stress magnification

The theory of elasticity had shown that defects or holes in infinite isotropic plates of linear elastic materials amplify stresses at their boundaries. They are **stress raisers**. The amount of stress magnification depends primarily on the shape, location, and orientation of the defect. In two dimensions, the two extremities of an elliptical hole lying orthogonal to the direction of the remote (e.g. regional) and uniaxial tensile stress σ are head-points where the boundary is parallel to stress. The amount of stress magnification at these two tips is related to both the pore geometry, i.e. the long (2L) and short (2ℓ) axes of the elliptical hole, and to the far-field stress. The maximum stress σ_{max} at the tips of the ellipse is given by:

$$\sigma_{max} \approx \sigma\left(1+\frac{2L}{\ell}\right)$$

where $\left(1+2L/\ell\right)$ is the **stress concentration factor**. Thus, the larger is the axial ratio, the greater the stress concentration. The ellipse flattens to a crack if $L >>> \ell$. For circular holes, $L = \ell$ and the stress at the two "extremity" points is three times larger than the applied stress $\left(\sigma_{max} = \sigma_3\right)$, independently of the hole's size.



The theory expresses stress concentration as a function of r, the radius of curvature of the ellipse at its ends, which is related to its length and width:

$$r = \ell^2/L$$

Solving this for ℓ and substituting into the L/ℓ ratio in the σ_{max} equation, the tip stress attains the form:

$$\sigma_{max} = \sigma\left(1 + 2\sqrt{L/r}\right)$$

and if $L >>> r$:

$$\sigma_{max} = 2\sigma\sqrt{L/r}$$

This criterion suffers from one major drawback: if $r \rightarrow 0$ (the case for a longitudinal crack), then $\sigma_{max} \rightarrow \infty$. This is not realistic because no material can withstand infinite tensile stress and there is no obvious tendency for cracks to lengthen spontaneously.

Exercise

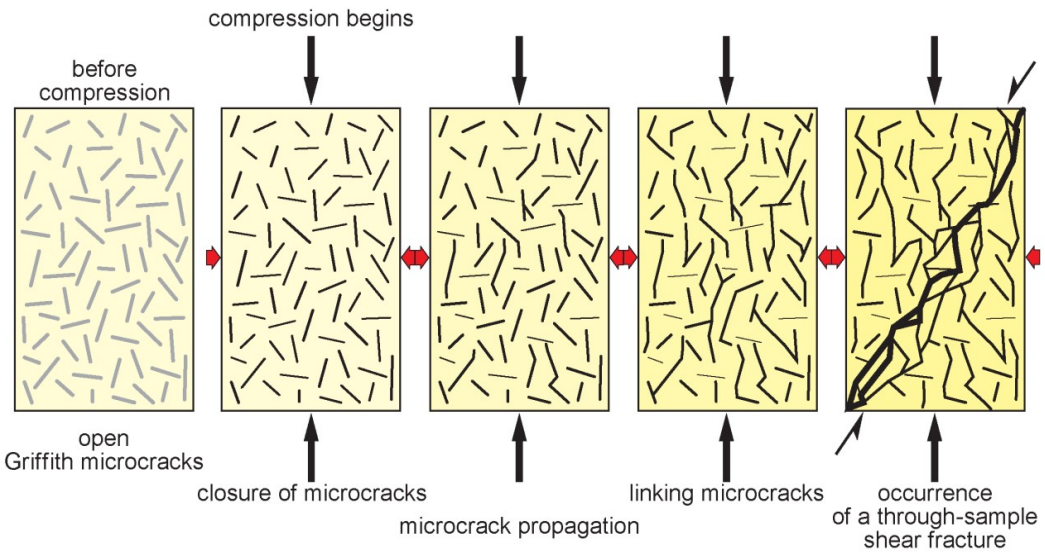
Numerically play with the shape of holes (from circle to flat ellipses) to visualize effects of porosity shape on stress concentration.

Postulate

Griffith postulated in 1920 that numerous sub-microscopic flaws such as bubbles, fine pores or simply grain boundaries, now called **Griffith cracks**, are naturally spread throughout apparently homogeneous and elastic materials. Griffith cracks in rocks may be original or induced imperfections along grain boundaries or within grains.

Concept

Such minute openings have the shape of extremely elongated and flattened ellipsoids across which atomic bonds are originally broken. A relatively low applied stress is variably amplified, depending on the orientation of the flattened ellipsoid with respect to the applied stress, at the sharply curved ends of these pre-existing cracks. Amplification produces stresses much higher than the mean stress in the material, to the point where the amplified local stress reaches the strength required to break the atomic bonds. Then the cracks spread spontaneously from their tips, where only a few chemical bonds have to be broken, at applied stress smaller than the theoretical tensile strength of the material. Progressive rupture and connection of the network of microcracks ultimately lead to brittle failure.



Development of a shear fracture in compression through coalescence of Griffith microcracks

Necessary condition for crack growth in two dimensions

Griffith's analysis integrated the elastic strain energy required to create the new surface area of an expanding single crack in a thin plate. His approach abides by the first law of thermodynamics: "*When a system goes from non-equilibrium to equilibrium state there is a net decrease in energy".* In a very small material volume, brittle fracture occurs when the decrease in the strain energy during an incremental crack growth is equal or exceeds the energy absorbed to create the new surfaces of the crack. The applied external load creates the elastic energy stored in the cracked plate. The thermodynamic argument is complex and can be found in textbooks dedicated to fracture mechanics. The demonstration shows that the tip tensile stress σ_T at either end of the crack equals the atomic bonds when:

$$\sigma_T = \sqrt{\frac{2AE}{\pi\lambda}} \tag{10}$$

where A is the surface energy per unit area of the crack (i.e. the energy required to create new surface), E is the effective Young's (elasticity) modulus of the rock and $\lambda = L/2$, i.e. the elliptical crack half-length. Assumptions imply that the crack extends in its plane. Equation (10) shows that short cracks require larger far-field stresses to grow than long cracks. In other words, the greater is the length of the cracks, the lower the macroscopic tensile strength of the material.

Experiment:

Take a piece of paper and pull strong on both extremities. How difficult is it to tear apart? Make a cut in the middle? Is it easier? Make on a similar piece of paper two cuts of different lengths. Result?

Critical stress intensity

Equation (10) is a result expressed in terms of far-field load. It can also be expressed in terms of the stress at the crack tips. The necessary thermodynamic criterion for crack propagation through opening only (i.e. motion orthogonal to L, with no lengthwise shear displacement) is:

$$K_I = \sigma_T \sqrt{\pi\lambda}$$

where K_I is the critical **stress intensity factor** also called **fracture toughness** at fracture propagation. It is a material property expressed in MPa.m$^{1/2}$. It defines fracture propagation in the considered material. K_I depends on parameters such as temperature, confining pressure and chemical/fluid environment. At room temperature, it varies from about 0.1 MPa m$^{1/2}$ for coal to 3.5 MPa m$^{1/2}$ for granite and dunite.

Theory of failure

Griffith extended the concept to materials containing a large number of elliptical cracks oriented at random. He assumed that these cracks propagate when the tensile stress reaches the critical stress intensity. In materials with cracks of different axial ratios, cracks with the largest axial ratios most likely grow first unstably and lengthwise into intact material. They interact and link under the influence of applied tensile or compressive stress to reach the failure of the rock. This theory leads to Griffith's criterion for failure expressed by a curved line in a Mohr diagram:

$$(\sigma_1 - \sigma_3)^2 = 8T_0(\sigma_1 + \sigma_3) \qquad \text{if} \qquad \sigma_1 + 3\sigma_3 \geq 0$$

and

$$\sigma_3 = -T_0 \qquad \text{if} \qquad \sigma_1 + 3\sigma_3 \leq 0$$

where T_0 is the uniaxial tensile strength of the material (the lowest intercept between the failure envelope and the horizontal, normal stress abscissa).

Note also that:

$\sigma_3 = 0$ then $\sigma_1 = 8T_0$

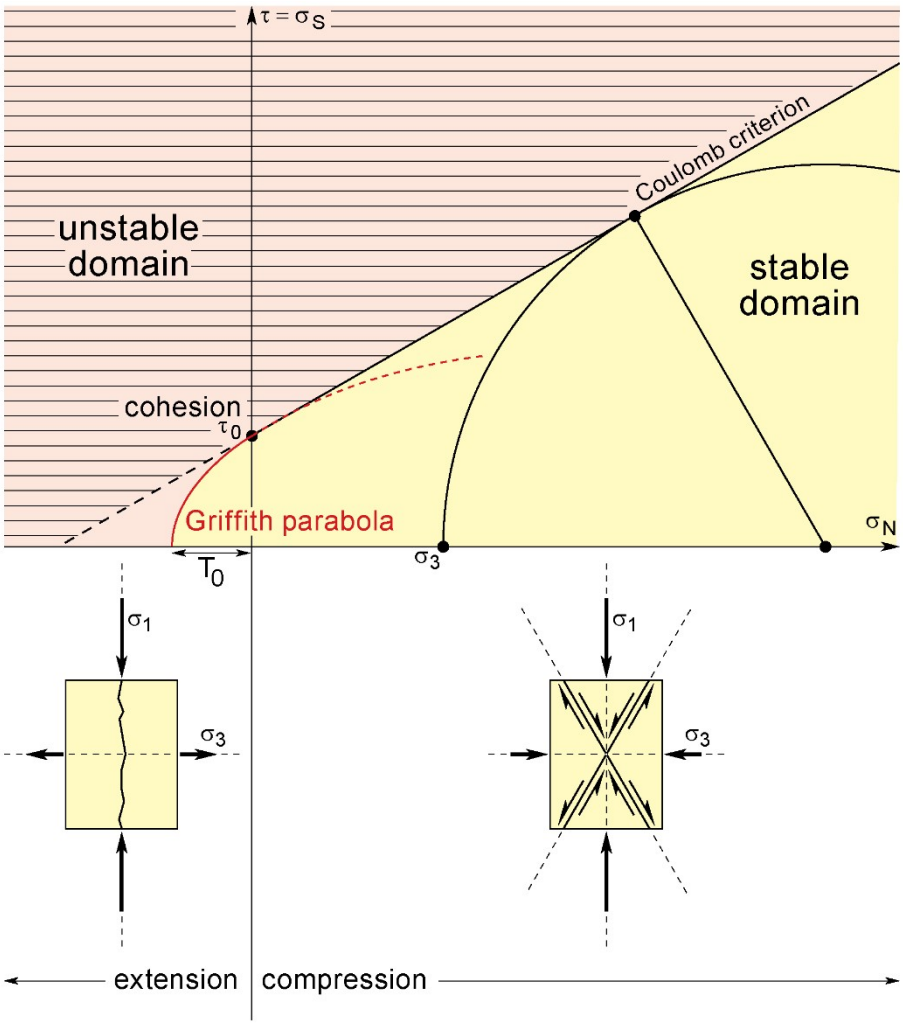
which suggests that uniaxial compressive stress at crack expansion is always eight times the uniaxial tensile strength. This is not consistent with experiments as the uniaxial compressive strength of most rocks is 10 to 50 times the uniaxial tensile strength. Typically, $T_0 \approx 40\text{MPa}$ for intact rock.

Graphical representation

The Griffith criterion can be expressed in function of normal stress σ_N and shear stress σ_S acting on the plane containing the major axis of the crack:

$$\sigma_S^2 = 4T_0(T_0 - \sigma_N) \tag{11}$$

This equation gives a parabolic envelope for failure points in the tension side of the plot.



Representation of the Griffith failure criterion
$\sigma_S^2 = 4T_0 (T_0 - \sigma_N)$
with respect to the Coulomb criterion in the two-dimensional Mohr diagram

On the compression side, equation (11) becomes:

$$\sigma_S = 2T_0 + \mu\sigma_N$$

which is a modified version of the Coulomb criterion (equation 6). When $\sigma_N = 0$, $\sigma_S = 2T_0$, which represents cohesion c . This relationship $(c = 2T_0)$ fits well experimentally derived curves for shear failure.

Brittle failure criterion

In experiments, initial failure occurs at the peak stress that the rock can sustain, known as **static friction**. The complete criterion for brittle failure is obtained by linking the two criteria (Griffith to Mohr-Coulomb; equations 6 and 10) at the point where they meet. For most rocks, the coefficient of sliding friction introduced to account for postulated closure of Griffith cracks under compression is $0.5 < \mu < 1$ (0.75 is a representative value). Linking the parabolic Griffith criterion to the straight Coulomb criterion explains a complete transition, with increasing mean stress (therefore depth), from extension fractures that are parallel to σ_1 (with negative σ values, left of the σ_S ordinate axis where $\sigma_N = 0$) to faults inclined at ca. 30° to σ_1 through shears with a dilatational component inclined at lower angles to σ_1. Experiments and field evidence corroborate the occurrence of shear fractures at very low angles to σ_1.

The type of fracture that occurs within intact rocks depends on the ratio of differential stress $(\sigma_1 - \sigma_3)$ to tensile strength T_0. For $\mu = 0.75$:

compressional shear failure occurs when $(\sigma_1 - \sigma_3) > 5.66T_0$,

extensional shear requires $5.66T_0 > (\sigma_1 - \sigma_3) > 4T_0$

extension fracturing requires $(\sigma_1 - \sigma_3) < 4T_0$.

Effects of environmental and material factors

Compressive strength is the capacity of a material to withstand axially directed compression. The compressive strength of rock is usually defined by the **ultimate stress**, the maximum stress the rock can withstand. Well-controlled compression tests have generated complete stress-strain curves of various rock specimens, so deciphering the role of various material and physical/chemical parameters.

Effects of fluids

Pores in rocks are primary features like vesicles in volcanic rocks and intergranular spaces in any rock. Secondary pore spaces are deformation-induced microcracks, joints, and faults. Pores of natural rocks contain fluids (e.g. water, oil, gaseous phase, and melt in deep levels of the Earth) which affect the failure of rocks in two ways: (1) the mechanical effect of fluid pressure, which accelerates the propagation of microfractures and (2) the chemical interactions between the rock and the fluid; this effect is attributed to the weakening of the crystalline framework at the highly stressed tips of cracks by a **stress corrosion** process involving rapid hydrolysis of silicon-oxygen bonds.

Fundamentally, fluid pressure counteracts and therefore decreases the confining pressure. When stress is applied to wet rocks, volume changes first translate into pore pressure changes. Excess pore pressure may or may not escape depending on the permeability of the rock at the time available.

Remember few definitions

Porosity describes how densely the material is packed. It is the ratio of the non-solid volume to the total volume of material. Porosity, therefore, is a dimensionless fraction between 0 and 1. The value ranges from <0.01 for granite, up to 0.5 for porous sandstone. It may also be represented in percent terms by multiplying the fraction by 100%. Porosity provides the void for fluids to flow through rock material. High porosity therefore naturally leads to high permeability.

Water content is a measure indicating the amount of water the rock material contains. It is the dimensionless ratio of the volume of water to the bulk volume of the rock material.

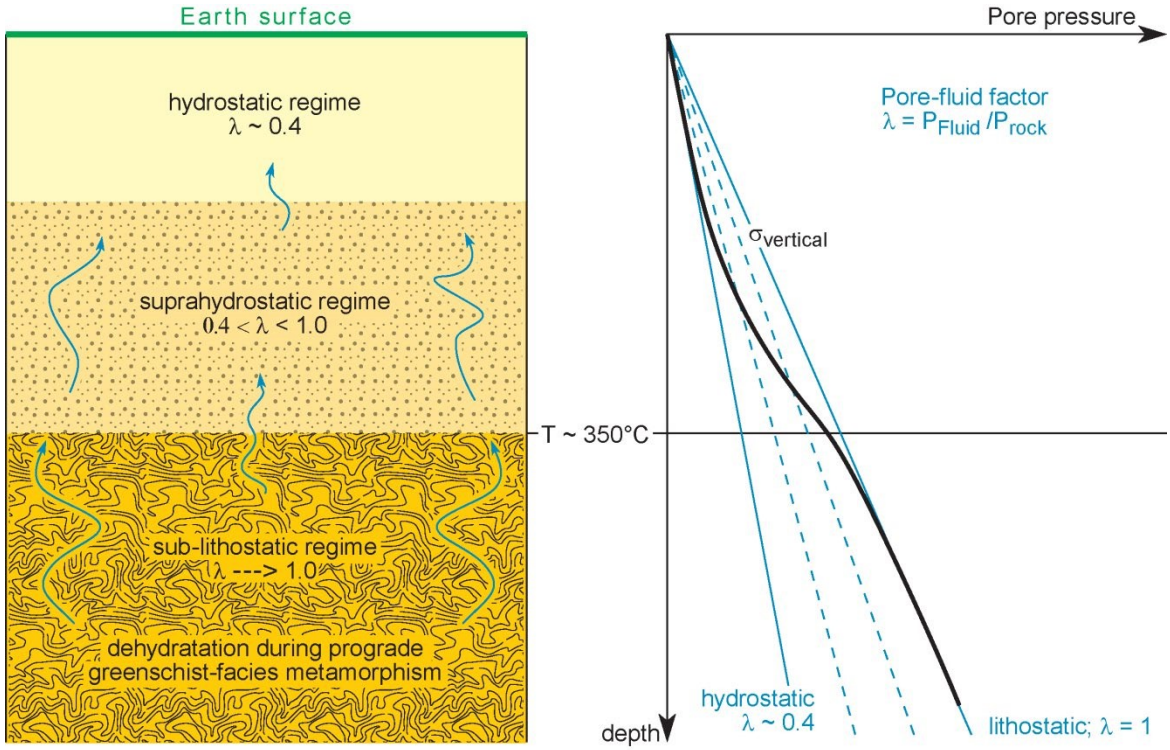
Permeability is a measure of the ability of a material to transmit fluids. Rocks generally have very low permeability. Permeability of the rock mass presents limited interest because the fluid flow is concentrated in fractures. The physical unit of permeability is m^2.

Pore pressure

If the pores are interconnected and communicate with the Earth's surface, which is common in the uppermost crust, the **hydrostatic pressure** in pore water at any depth is equal in all directions, outward from the pore space, to the weight of the water column from this depth to the surface.
Therefore, the pore fluid pressure within a rock increases as the rock is buried. To quantify this consideration, the hydrostatic pore pressure at any depth is about 0.25 - 0.3 times the lithostatic or rock pressure, if one takes the mean density of the pore fluid to be 1.0 g cm^{-3} and the mean density of the rock column to be 2.5 to 3.0 g cm^{-3}. However, the pore pressure can vary temporarily and spatially in the lithosphere in a wide range of values.
The **pore-fluid factor** λ is the ratio between the fluid pressure and the lithostatic pressure:

$$\lambda = P_{fluid}/P_{rock}$$

which defines the fluid-pressure level at different depths.



Hypothetical profile of pore pressure in the upper crust
after Sibson (2004) *J. Struct. Geol.* **26**, 1127-1136

For dry rocks $\lambda = 0$.
Ratios of pore pressure to lithostatic pressure above 0.8 have been measured in oil fields. Ratios approaching 1.0 are considered for water-saturated, unconsolidated fresh sediments.
For hydrostatically pressured sediments $\lambda = 0.4$
If pore fluids are not connected to the surface, several mechanisms can generate **fluid overpressure** (pore pressures greater than hydrostatic, i.e. $\lambda > 0.4$). Important mechanisms are:

- Seismic shocks can rapidly amplify the fluid pressure for a short time, which may even fluidize water-saturated soil or unconsolidated sediments.
- Reduction of pore space (compaction) in fluid-rich sediments by burial or tectonic deformation and rapid fluid release by dehydration of mineral assemblages during diagenesis. Rates of loading or dehydration may exceed the escape rate of pore fluids in low permeability rocks. This is particularly true if pores do not form an open system connected with the surface.
- Deeper in the crust fluids internally expand at elevated temperatures and new fluids are added by fluid-releasing metamorphic reactions.

- Magmatic intrusions and melt segregation in melting rocks generate fluids faster than they can escape.

Under these circumstances, the presence of fluids can considerably reduce the effective brittle strength of rocks. A rock overpressured by fluids in great depths may break as if it were near the Earth's surface. Fluid-triggered crack propagation likely assists emplacement of magmatic intrusions, in that of dikes and sills.

Effective normal stress

The stress state within the pores is hydrostatic and the fluid pressure acts so as to oppose the lithostatic stress caused by the overburden: Pore fluids support some of the load that would otherwise be supported by the rock matrix and, therefore, the pore pressure P_f is a component of the total stress.

P_f acts uniformly against the normal stress σ_N on the rock according to the equation:

$$\sigma_{eff} = \sigma_N - P_f$$

The pore pressure reduces all the lithostatic stresses by an amount P_f to give the **effective normal stress** σ_{eff}. Thus the principal stresses become:

$$\sigma_1^{eff} = \sigma_1 - P_f$$
$$\sigma_2^{eff} = \sigma_2 - P_f$$
$$\sigma_3^{eff} = \sigma_3 - P_f$$

The pore pressure produces no shear stress, and hence no shear deformation. The net result of pore fluid pressure is to allow the rock to behave as if the confining pressures were lowered by an amount equal to P_f.

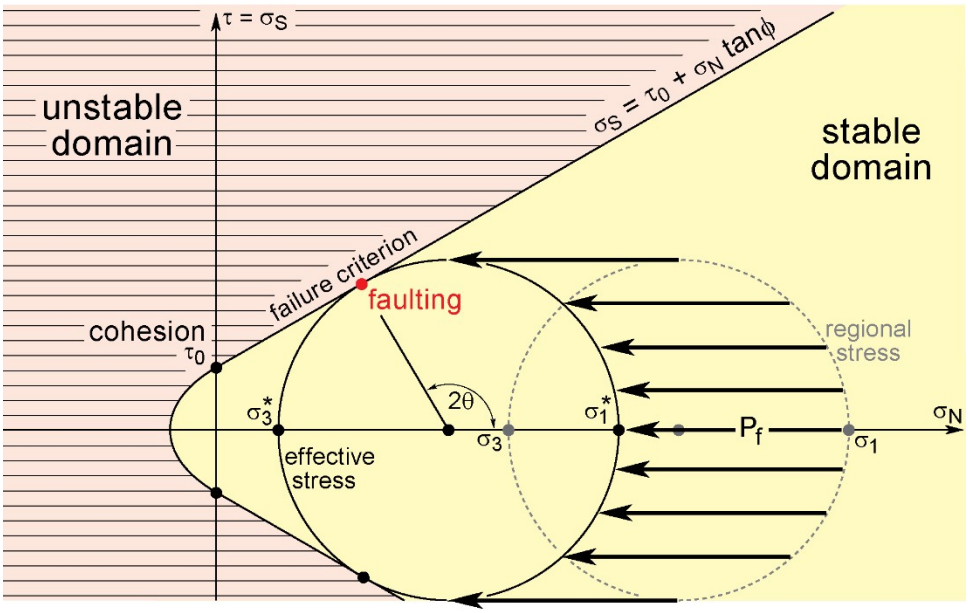
Graphical representation

A Mohr diagram can illustrate how pore pressure promotes faulting.

- A circle representing the state of stress in a rock with a zero pore pressure is drawn. The effective normal stress σ_{ii}^{eff} equals the total normal stress σ_{ii}. The rock is stable under these conditions if the circle is below the failure envelope.
- If the pore pressure is increased gradually (for example during burial), σ_1^{eff} and σ_3^{eff} are smaller than σ_1 and σ_3, respectively, but the differential stress does not change:

$$\left(\sigma_1^{eff} - \sigma_3^{eff}\right) = \left(\sigma_1 - \sigma_3\right)$$

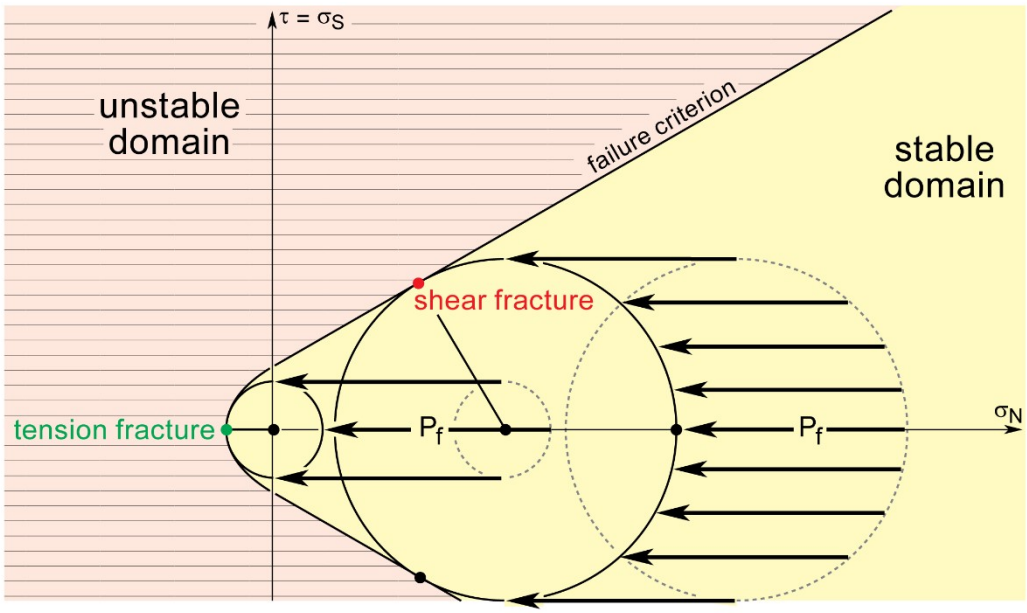
  The 'dry' circle representing the regional stress keeps the same size but inexorably migrates along the normal stress abscissa from its original position towards lower compressive stresses, closer to the failure envelope. The amount of translation of the stress circle is determined by the magnitude P_f of the fluid pressure.
- If P_f is large enough, the circle will hit the Mohr envelope and faulting will occur. The pressure of the pore fluid thus allows faulting even though the shear stresses are too small for faulting in the dry rock, or wet rock at lower pore pressures. This effect is verified in triaxial tests. Internal fluid pressure reduces markedly both the fracture strength and the ductility of rocks, which both are functions of **the effective confining pressure**. This effect also accounts for the increased occurrence of landslides in the aftermath of heavy rainfall.



Effect of a pore pressure P_f represented in a Mohr diagram

Hydraulic fracturing

By simply increasing the fluid pressure, the outward push of the fluid creates tensile stress sufficient to cause crack propagation at the pore and crack tips. This process is called **hydraulic fracturing**. In this way, an originally compressional stress regime can be changed so that one or more of the principal stresses becomes effectively tensile and the conditions for tensile failure can be satisfied.



Effect of the differential stress on the fracture type in two-dimensional Mohr diagram

A fluid opens a fracture if the fluid pressure P_f equals or exceeds the normal stress σ_N acting on the fracture. A vertical line locating the fluid pressure divides the Mohr circle into two fields:

- A domain with $P_f \geq \sigma_N$ (left side of the fluid pressure line) in which fractures can dilate.
- A domain with $P_f \leq \sigma_N$ (right side of the fluid pressure line) in which fractures remain closed.

The condition for fracture opening is expressed from equation (3) as:

$$P_f \geq \frac{\sigma_1 + \sigma_3}{2} + \frac{\sigma_1 - \sigma_3}{2} \cos 2\theta$$

The **driving stress ratio** D_{SR} defines the ranges of orientations of fractures that can dilate under a given fluid pressure:

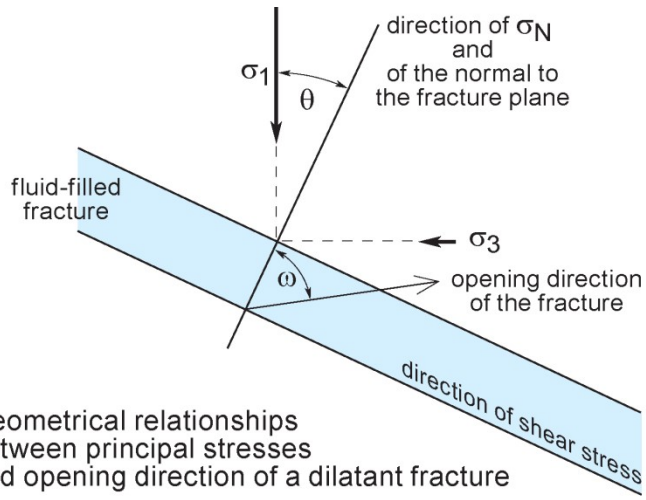
$$D_{SR} = \frac{P_f - \frac{\sigma_1 + \sigma_3}{2}}{\frac{\sigma_1 - \sigma_3}{2}} \geq \cos 2\theta$$

which is expressed more simply in terms of mean stress $\overline{\sigma}$ and maximum shear stress $\sigma_{S\,max}$ defined in the lecture on stresses:

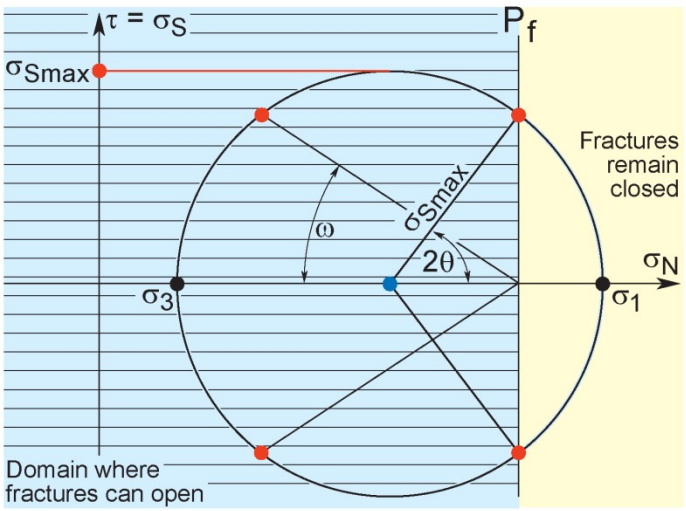
$$D_{SR} = \frac{P_f - \overline{\sigma}}{\sigma_{S\,max}} \geq \cos 2\theta$$

A Mohr diagram represents the range of fracture orientations that will open under this bounding condition. The pore pressure, the intersection of the Mohr circle with the vertical line passing through the pore pressure, and the center of the Mohr circle (the mean stress) define a triangle whose horizontal side, lying on the normal stress axis, is $P_f - \overline{\sigma} = \sigma_{S\,max} \cos 2\theta$. Fracturing represented by the intersection point on the Mohr circle is thus written:

$$P_f - \overline{\sigma} \geq \sigma_{S\,max} \cos 2\theta$$



Geometrical relationships between principal stresses and opening direction of a dilatant fracture



Delimitation of fracture dilation domains in function of pore fluid pressure P_f in two-dimensional Mohr diagram

This relationship is an expression of D_{SR}, which uses the relative magnitudes of stresses and fluid pressure to predict the range of fracture orientations able to dilate.
From the Mohr construction one readily sees that:

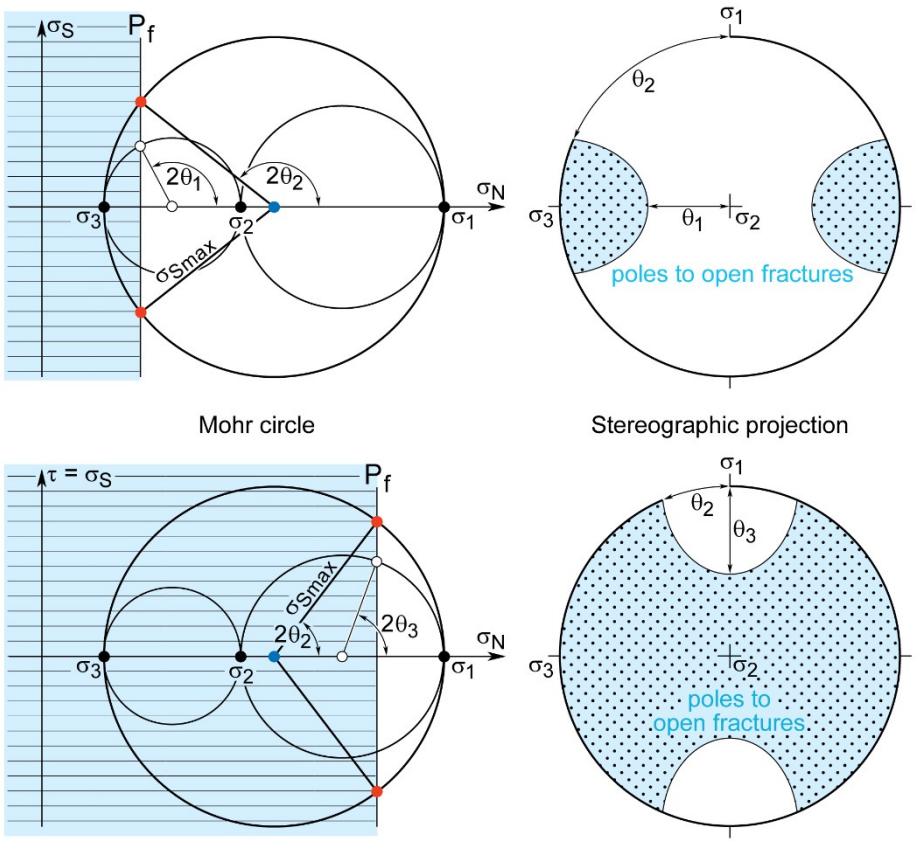
If $D_{SR} < -1$, $P_f < \sigma_3$, no fractures open.

If $-1 < D_{SR} < 1$, $\sigma_3 < P_f < \sigma_1$, a limited range of fracture orientations can open.

If $D_{SR} > 1$, $P_f > \sigma_1$, fractures of any orientation can open (brecciation of the rock).

The same reasoning and the same constructions can be used in three dimensions.
When $P_f < \sigma_2$, the poles to fractures able to dilate define a cluster distribution around σ_3. A narrow cluster region indicates a low driving fluid pressure ($P_f \approx \sigma_3$).
When $P_f > \sigma_2$, poles to fractures able to dilate form a girdle distribution perpendicular to σ_1. A wide pole distribution indicates a high driving fluid pressure ($P_f \approx \sigma_1$).

Accordingly, the structural expression of hydraulic fracturing can vary from randomly oriented extensional fractures (resulting in a breccia when $P_f > \sigma_1$) through aligned extensional fractures to shear fractures.



Geometrical relationship between poles to open fractures and principal stress axes when the fluid pressure P_f is less than σ_2 (top) and greater than σ_2 (below)
after Jolly & Sanderson (1997) *J. Struct. Geol.* **19(6)**, 887-892

Note that this mechanism is independent of the depth at which it may occur. This process is routinely applied in the petroleum industry to create fractures in low-permeability rocks. Besides, crack propagation increases space for fluid whose pressure diminishes unless there is additional fluid entering the system. Since crack propagation takes place only at some yield strength, hydraulic fracturing typically occurs in pulses, when pore pressure reaches the necessary value.
Note also that tensile failure can occur in rocks without the aid of a high internal fluid pressure: for example, during the contraction of a layer as a result of desiccation of sediment or the cooling of an igneous body.

Fluid pressure in shearing resistance

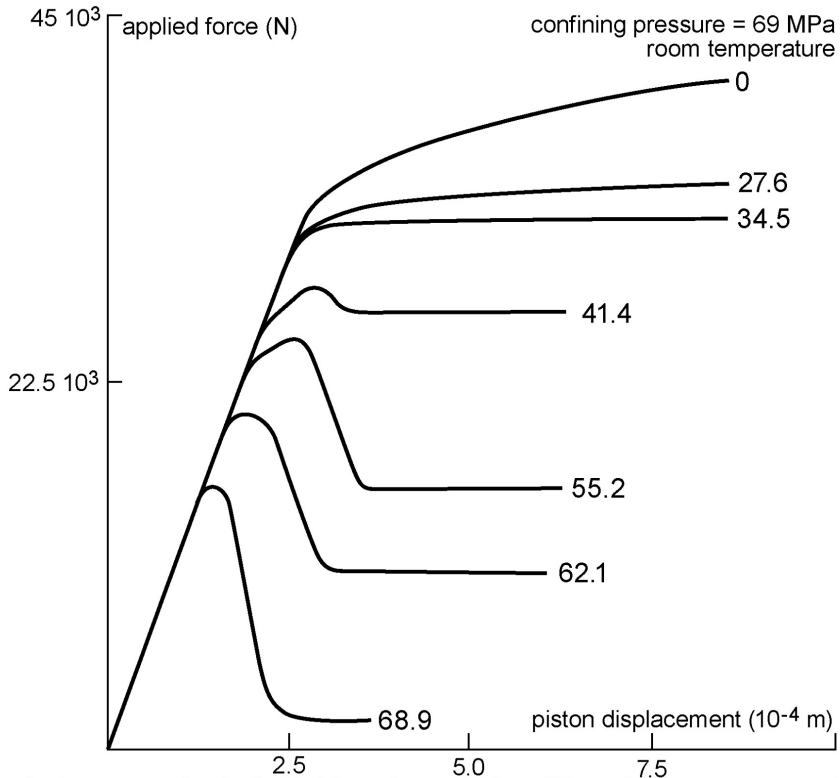
The **shearing resistance** of the rock is the shear stress on a potential fault plane that is just sufficient to initiate fault movement. The conditions for sliding along a plane in fluid-saturated materials are given by a simple modification of the Mohr-Coulomb criterion for failure (equation 6),

$$\sigma_S = c + \mu(\sigma_N - P_f) \tag{12}$$

This equation has several inferences.

- A rock will exhibit essentially the same shearing resistance when σ_N = 1 kbar, $P_f = 0$ and when σ_N = 2 kbar, $P_f = 1$ kbar, because the effective normal stress is 1 kbar in both cases. The effect discussed here is one in which increased fluid pressure changes the state of stress throughout the rock, which somehow reduces the shear stress required for faulting.
- While the normal stress σ_N tends to strengthen the fault plane by pushing together the opposing rock blocks, hence increasing friction, the fluid pressure acts to weaken the fault by pushing the opposing rock blocks apart. In other words, increased pore pressure dampens down the effect of confining pressure, which strengthens a fault as it increases with depth.

In both cases, pore pressure can be thought of as having a lubricating effect on faults in the sense that it reduces frictional, shearing resistance to movement.



Stress-strain curves for Indiana Limestone under different pore pressures in MPa
from Robinson, 1959, *Quart. Colo. School Mines*, **54**, 177-199

Faults with favorable orientations for slip or dilation present potential fluid flow pathways. Therefore, recent research on earthquake control is much concerned with pore pressure effects because it is the one critical variable that can be manipulated to some extent by man. The basic idea is that a suitable local increase in pore pressure, brought about by pumping fluids into drill holes, could lower the shearing resistance of rocks sufficiently to trigger local faulting and small earthquakes. Many such small earthquakes may gradually release the stored energy that would otherwise accumulate for a single large and catastrophic earthquake. A correlation between the rate of fluid injection and

earthquake frequency has been observed. These studies tend to confirm that small earthquakes can be triggered or suppressed by judicious control of pore pressure, but it is not yet clear what effect such control of small earthquakes will have on the occurrence of large ones.
High pore pressures are considered to play a vital role in some examples of low-angle thrust faulting. Thrust sheets 30-100 km wide pose serious mechanical problems if they moved dry because friction along their bases would seem to require either impossibly high shearing stresses in the sheets (if they are pushed from behind) or unreasonably steep slopes (if they slide downhill under the influence of gravity). Pore pressures higher than normal could allow thrust sheets to be pushed more easily or to slide down slopes with dips as low as a degree or two. However, an alternative explanation for large thrust sheets is that the basal shear zone, or detachment horizon, dominantly behaved as a viscous material and not as a frictional material.
Equation (12), like the Coulomb equation for dry rocks (6), fits many experimental data. However, this equation relates the shearing resistance entirely to the pressure of the pore fluid when chemical properties of the fluid may also be important, for example, in controlling the rate of stress corrosion at the tips of cracks.

Mode of deformation

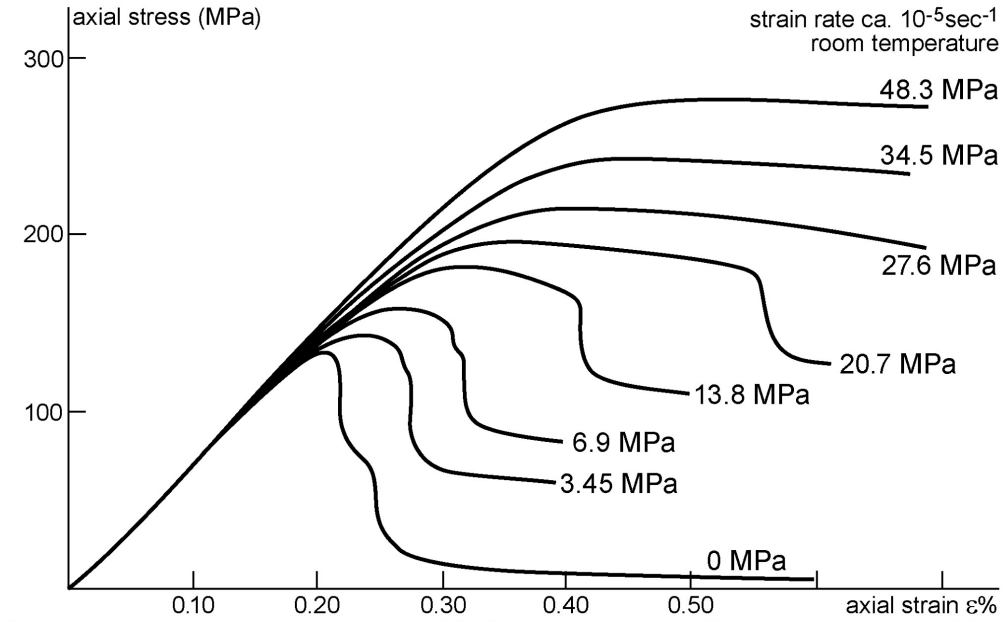
A series of triaxial compression tests with a constant confining pressure but with various pore pressures illustrates the influence of pore-water pressure on the behavior of porous rock. There is a transition from ductile to brittle behavior as pore pressure is increased from 0 to higher values. Fluid pressure lowers the elastic limit, thus increases the ductility field of rocks (i.e. their capacity for change of shape without gross fracturing) in experiments. In that way, pore pressure influences the mode of deformation, for example from ductile under moderate confining and fluid pressures to brittle under similar confining pressure but high pore pressure. The current hypothesis is that when the frictional resistance is higher than the shearing strength of rock, the rock shows ductile behavior. However, this brittle/ductile transition is still incompletely understood.

Dry rock

If the pore pressure can dissipate, it tends to zero. The rock is most likely work-hardened.

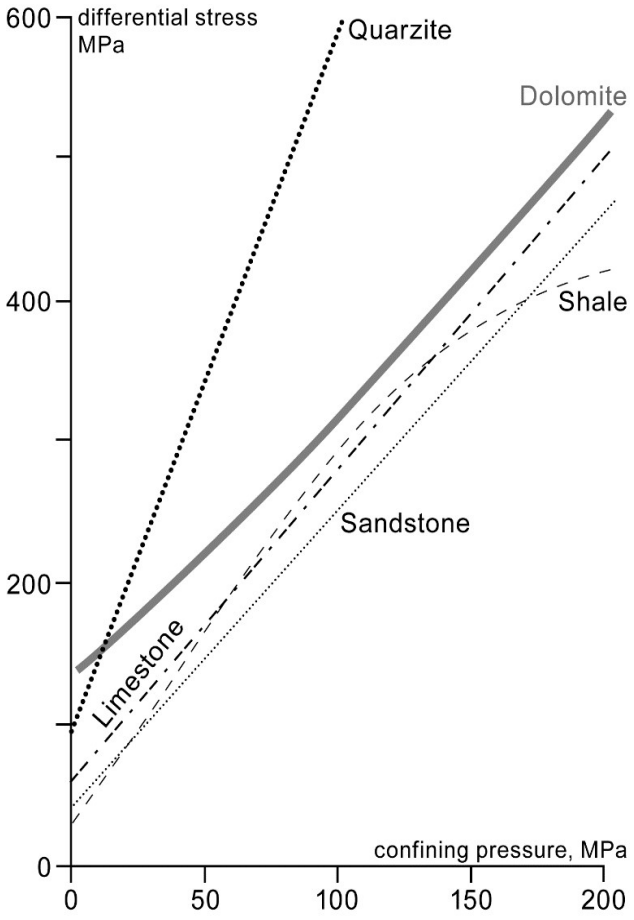
Effects of confining pressure

Triaxial compression tests at various confining pressures show that with increasing confining pressure:
- the peak strength increases;
- there is a transition from typically brittle to fully ductile behavior; the confining pressure that causes the post-peak reduction in strength is called the brittle-ductile transition pressure. The brittle-ductile transition pressure (ca 50 MPa for marble) varies with rock type;
- the region incorporating the peak of the axial stress-axial strain curve flattens and widens;
- the post-peak drop in stress to the residual strength reduces and disappears at high confining stress.
- The ductility of most rocks (i.e. how much strain is accommodated before reaching ultimate strength and failure) increases with increasing confining pressure. In the ductile regime, two lines parallel to the σ_N-axis in a Mohr diagram represent by the **Von Mises criterion**.



Stress-strain curves for Tennessee Marble at various confining pressure conditions
from Wawersik & Fairhurst 1970 *Int. J. Rock Mech. Min. Sci.* **7(5)**, 561-575

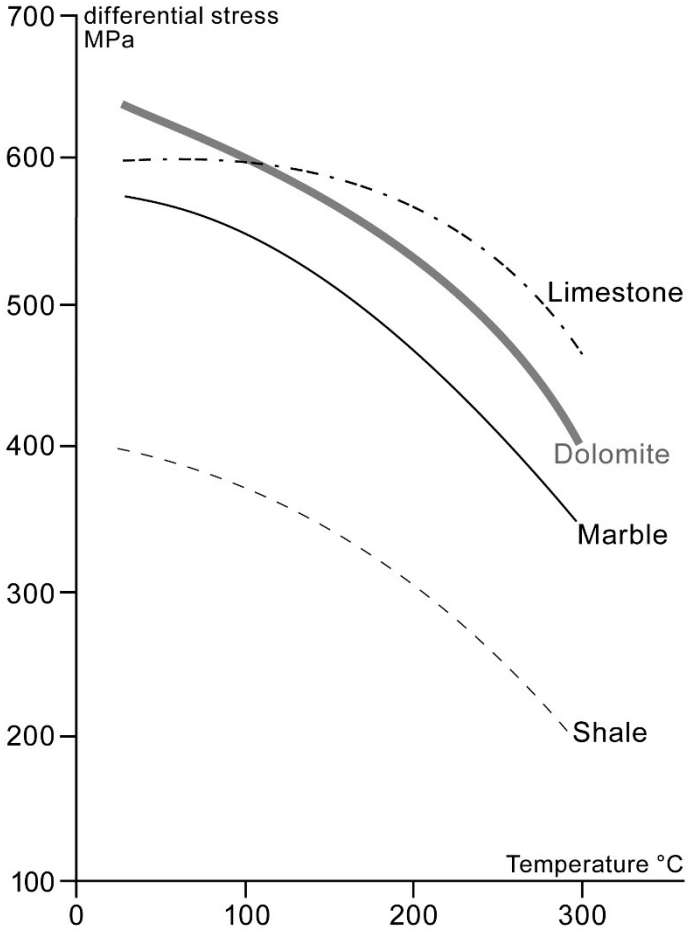
The dependence of shearing resistance on normal stress is shown by a differential stress / confining pressure diagram on which yield stress (arbitrarily taken as the differential stress at 2% strain) and ultimate strength of rocks is plotted. For most rocks both the yield stress and the ultimate strength increase almost linearly with increased confining pressure. In other words, rocks are stronger in confined compression and an increase in pressure suppresses the formation of new fractures.



Effect of confining pressure on shearing resistance
after Handin & Hager 1957 *Am.Ass.Petr.Geol.Bull.* **41**, 1-50

Effect of temperature

Temperature affects the mechanical properties of rocks much less than confining and pore pressures. From above 200 to 500°C, rocks are ductile. The yield stress and ultimate strength are reduced with increased temperature, but this reduction varies from one rock to another, depending on their mineralogical composition. Heating usually increases ductility.



Effect of temperature on shearing resistance
after Handin & Hager 1958 *Am.Ass.Petr.Geol.Bull.* **42**, 2892-2934

Effect of strain rate

Temperature affects the mechanical properties of rocks much less than confining and pore pressures, and effects are rather limited from above 200 to 500°C, according to their mineralogical composition, rocks are ductile. The yield stress and ultimate strength are reduced with increased temperature, but this reduction varies from one rock to another. Heating usually increases ductility.

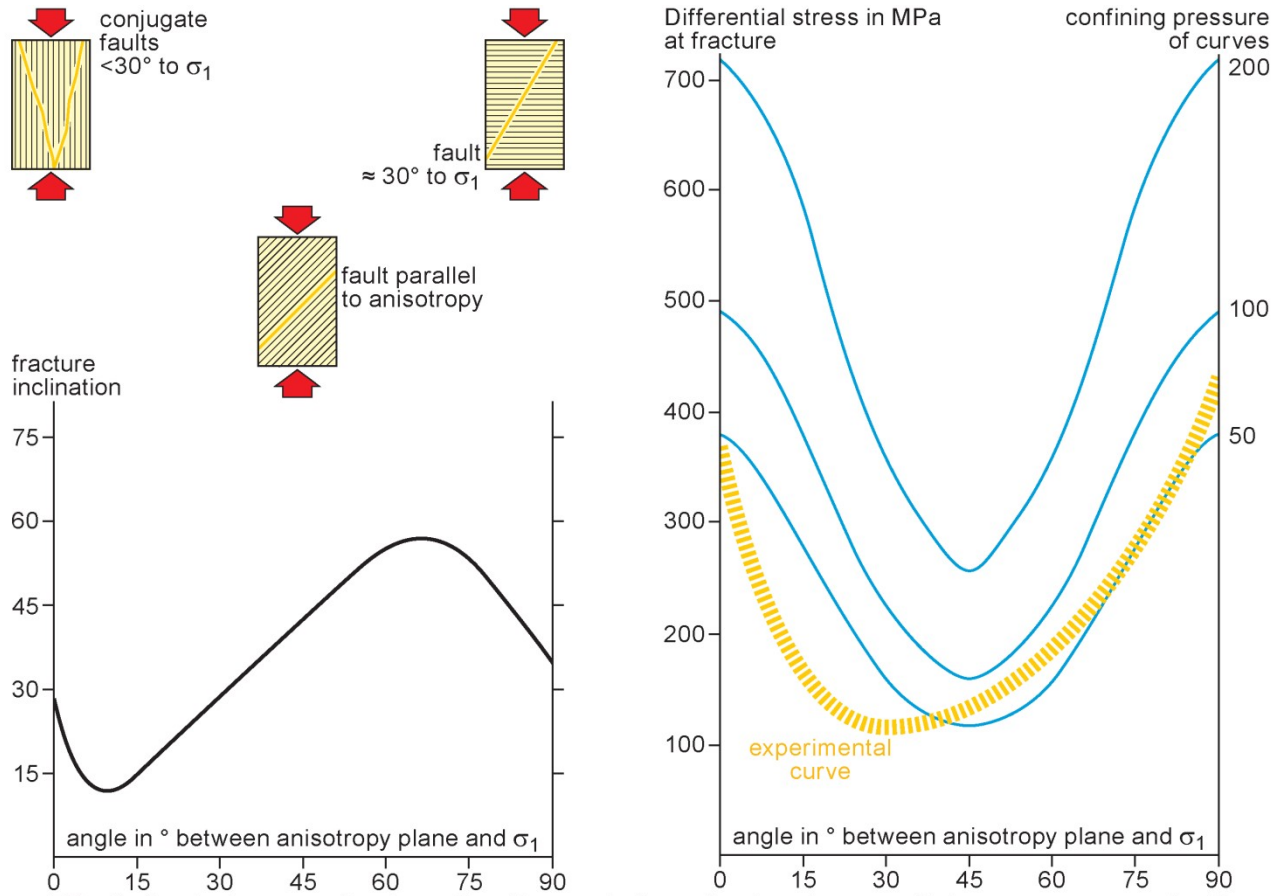
Effect of planar anisotropy

Rocks may contain well-developed planar anisotropy such as bedding and foliation. The intact rock material may be stronger in the direction of maximum shear stress than along the anisotropy. In this case, shear fractures are created along the planar anisotropy, therefore at angles that differ from those estimated from the σ_N/σ_S Mohr envelope. New shear fractures are forced by the orientation of the plane of anisotropy whose lower shear resistance reduces the shear strength of the bulk material. This textural weakening is a shear softening mechanism.

Peak strength

The peak strengths vary with the orientation of the plane of anisotropy (of weakness) with respect to the principal stress directions. Experimental curves of differential stress at failure versus inclination

of anisotropy are close to concave upward parabolas. The curves are lifted upward with increasing confining pressure (increase in breaking strength) while shear fractures tend to develop at small angles to the compression direction.



Effect of anisotropy on fracture direction and strength at various confining pressures and with respect to cleavage orientation of Martinsburg slates (after Donath 1972 *Geol. Soc. Am. Mem.* **135**, 95-128)

The analytical solution shows that the differential stress can be given by the equation:

$$\sigma_1 - \sigma_3 = \frac{2\left(c_w + \sigma_3 \tan\phi_w\right)}{\left(1 - \tan\phi_w \cot\beta\right)\sin 2\beta}$$

where c_w = cohesion of the plane of anisotropy

ϕ_w = angle of friction of the plane of anisotropy

β = inclination of the anisotropy plane to the compression direction

The minimum strength occurs when:

$$\tan 2\beta = -\cot\phi_w \qquad \text{or} \qquad \beta = 45° + \left(\phi_w/2\right)$$

This shows that when the rock containing an existing weakness plane that is about to become a failure plane, the rock has the lowest strength.

Orientation of shear fractures

In anisotropic rocks, the angle of faulting varies considerably, depending on the orientation of the anisotropy with respect to the principal stress direction.

- Rocks compressed parallel to the anisotropy sustain the greatest differential stress and shear fractures occur at less than 30° to the compression direction.
- Rocks compressed at 45 – 30 ° to anisotropy show the least strength.
- Shear fractures develop parallel to the anisotropy plane when the latter is inclined 15 to 45° to the compression direction.

- Rocks compressed orthogonal to the anisotropy sustain the greatest differential stress and shear fractures occur at 30° to the compression direction.

Slip tendency

Slip is likely to occur on a fracture when the resolved shear stress σ_S on that plane equals or exceeds the frictional resistance to sliding. Frictional resistance is proportional to the effective normal stress σ_{eN} acting across that surface. The slip tendency T_s of a surface is the ratio of maximum resolved shear stress to normal stress acting on that surface:

$$T_s = \sigma_S / \sigma_N$$

T_s is a measure of the relative likelihood that a fault or fracture will undergo slip. It is sensitive to both the form of the stress tensor and the orientation of the studied surface. The maximum value of T_s is limited by the slope of the sliding envelope. Whether or not a surface slips depends upon details of local conditions such as rock or fault cohesive strength, if any, the coefficient of internal friction and the orientation of the fault or fracture surface. T_s that causes slip on a cohesionless surface is often referred to as the fault strength in earthquake focal mechanism analysis. Under most crustal conditions, faults with $T_s \geq 0.6$ are ideally oriented for slip (Byerlee law).

Dilation tendency

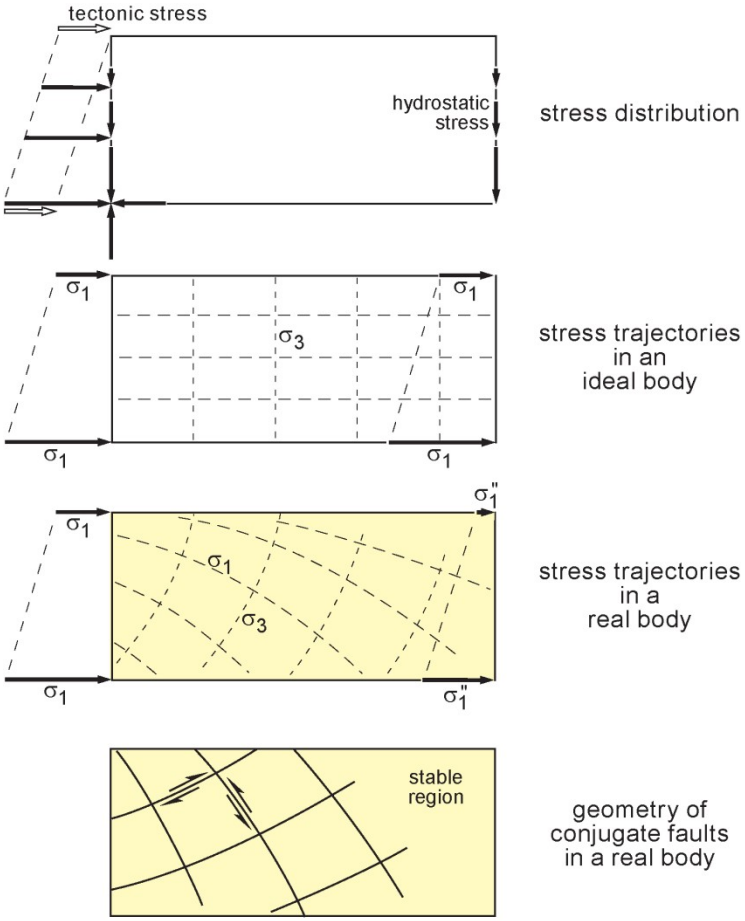
Dilation of fractures is largely controlled by the resolved shear stress, which is a function of lithostatic and tectonic stresses and fluid pressure. The normal stress on a fracture depends on the magnitude and direction of the principal stresses relative to the fracture plane. The ability of a fracture to dilate and transmit fluid is directly related to its aperture, which in turn is a function of the effective normal stress acting upon it. The normal stress can be computed for surfaces of all orientations within a known or hypothesized stress field. This normal stress can be normalized by comparison with differential stress. The resulting dilation tendency T_d for a surface is then defined as:

$$T_d = (\sigma_1 - \sigma_N) / (\sigma_1 - \sigma_3)$$

Maximum depth for faulting in dry rocks - listric faults

The dependence of shearing resistance on normal stress leads to the prediction that the resistance of dry rock to faulting should be greater at greater depths. This leads to the maximum depth for faulting if the shear strength of a given rock type becomes so high that deformation can occur by other mechanisms at lower stresses. In effect, the stress conditions required for flow by crystal slip are a limiting factor verified experimentally.

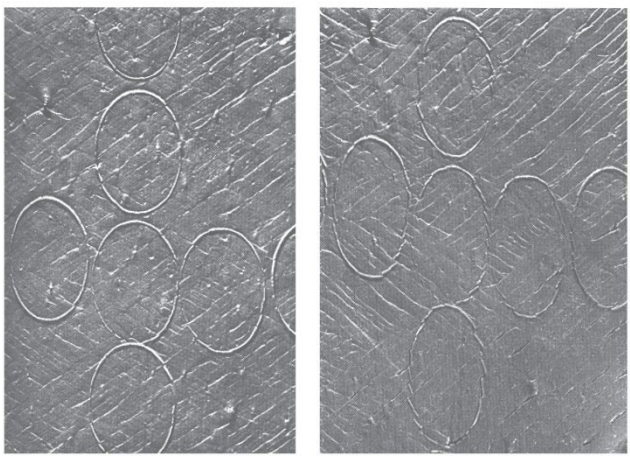
Note also that curved faults may also result from a uniform stress orientation due to the effect of downward increase in confining pressure. This change is shown again by the change in slope of the Mohr failure envelope. Ductile flow dominates at 10-20km depth, depending on rock composition and regional thermal gradients.



from Hafner (1951) *Bull. Geol. Soc. Am.* **62**, 373-398

Fault orientation relative to principal strain axes

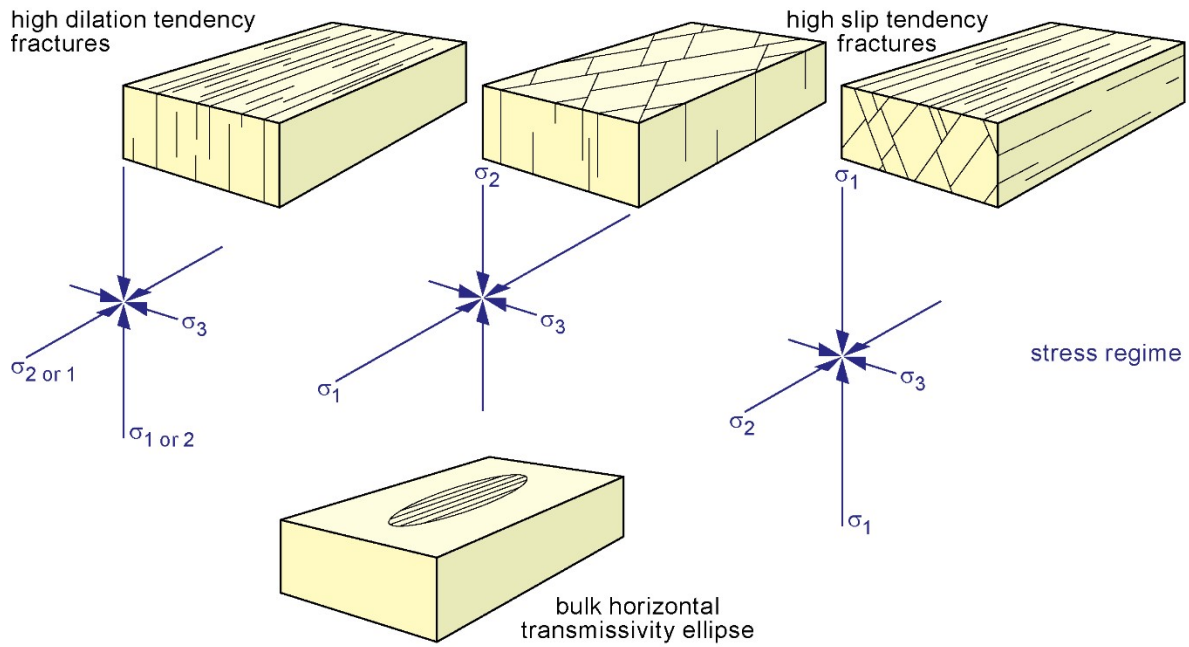
Conjugate faults, whether they are thrust, and normal or strike-slip faults, occur under plane strain conditions and intersect in a line parallel to the intermediate principal strain axis λ_2. The conjugate angle is normally an acute angle except where modified by internal rotations. The minimum principal strain axis λ_3 bisects the conjugate angle and occupies, along with the maximum principal strain axis λ_1, a plane perpendicular to λ_2. Minor structures like gashes and striations can be used to help to define the principal strain axes. Tension gashes are oriented perpendicular to the axis of maximum extension λ_1. Striations are parallel to the line of intersection of each fault with the $(\lambda_2;\lambda_3)$ plane.



Faulting in affine deformation
from Hoeppener *et al.* 1970 *Geol. Rundsch.* **59**, 179-193

Fault orientation and anisotropic porosity

Anisotropic permeability in fractured aquifers arises from the abundance and distribution of subparallel faults and fractures and permeability of associated damaged zones (fault breccias).



Effects of faults with high slip tendency or high dilation tendency on the development of anisotropic porosity in areas where the minimum principal stress σ_3 is horizontal
(adapted from Ferrill *et al.* 1999, *GSA today* **9(5)**, 1-7)

Faults and fractures orthogonal to the maximum principal stress tend to close, thereby reducing permeability perpendicular to the regional compression direction. Conversely, faults and fractures striking perpendicular to the minimum principal stress direction (i.e. parallel to the regional compression) tend to open, which relatively enhances their permeability. Regional stress combined with existing fractures thus influence the regional permeability and fractures favorably oriented for slip or dilation in the ambient stress field tend to be the most active groundwater flow pathways. This effect produces a transmissivity anisotropy.

Faulting History

Microfracturing and Dilatancy

When a rock is compressed and shortened at a constant rate in the laboratory, brittle deformation begins before any fault forms across the sample.

Up to a stress difference of roughly half the fracture strength, the specimen shortens by a fraction of 1 % and its volume decreases by an even smaller amount, governed by the compressibility of constituent minerals. Through this stage, the original dimensions of the specimen are recoverable and the behavior is elastic.

At stress differences greater than about half the fracture strength, inelastic effects begin to be noticeable. The volume of the specimen no longer decreases with continuing longitudinal shortening but instead increases slightly; the specimen is now **dilatant**. Acoustic emissions monitored during experiments and swarms of microseisms signal the opening of cracks and microfractures (the Griffith cracks) with limited propagation on a granular scale. The intensity of microfracturing activity increases while approaching the fracture strength. Microfracturing events become more frequent and spatially more concentrated near the eventual fault plane. Such observations suggest that

microfracturing and dilatancy may also precede fault displacements in nature. Shallow depth, dilatant faults thus produce open spaces in which secondary deposits may crystallize.

Lockner *et al.* 1991

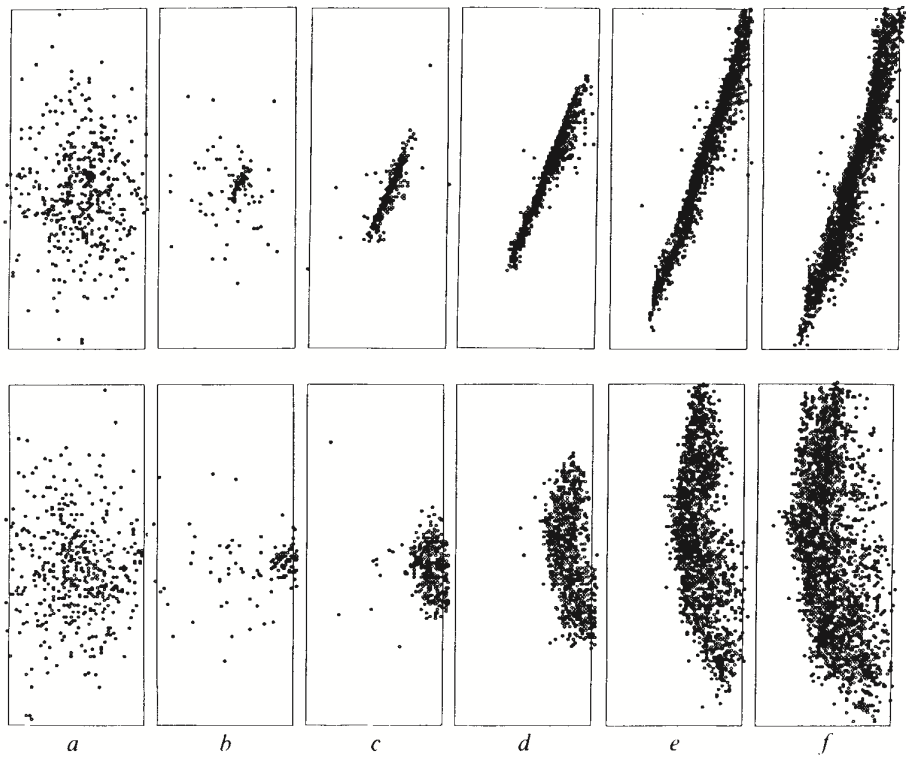


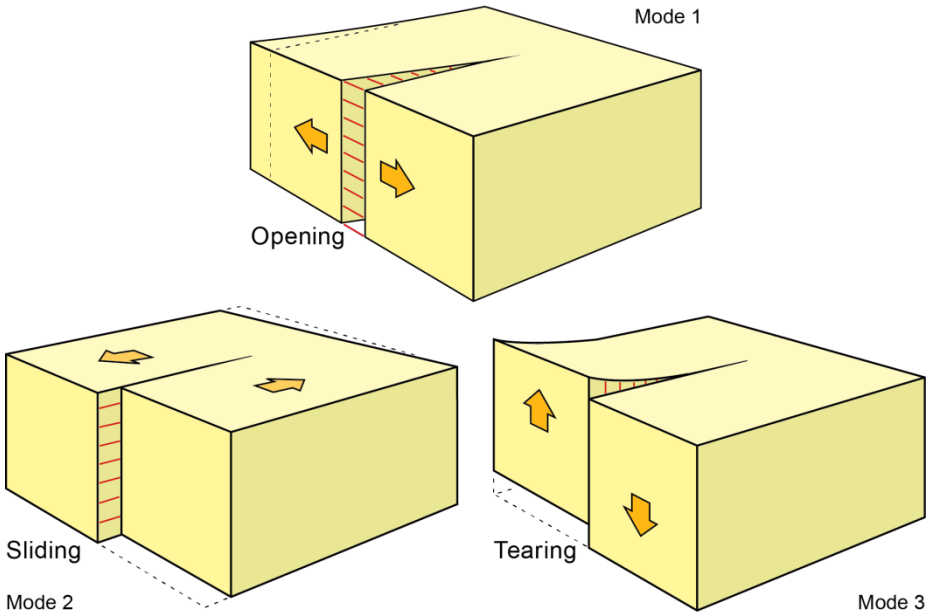
FIG. 2 Sequential plots of locations of acoustic emission. Stress interval for each plot is shown in Fig. 1. Upper plots show events viewed along-strike of eventual fault plane (seen as diagonal band of events). Lower plots show same events when fault plane is viewed face-on. Fault nucleates in *b* and propagates across sample in *c–f*. A distinct fracture front develops, in *d–f*, as fault grows. Number of events per plot in *a–f* are 474, 123, 402, 1088, 2292 and 4038, respectively.

According to Griffith's theory, a crack properly oriented with respect to the directions of the principal stresses propagates in the direction normal to the maximum tensile stress. The rate at which microfractures grow is an important factor determining the rate or stress difference at which brittle yielding will occur on larger scales. How propagating microfractures link up to give a macroscopic fault is still poorly understood, even in rocks deformed under laboratory conditions. In some tests, for example, extension microfractures form early and link up through suitably oriented grain boundary cracks to form a fault. In other tests, however, through-going faults are thought to predate the associated extension microfractures.

Fractures propagate from initial crack tips in three, or in a combination of the three mutually orthogonal ways. These three basic modes of displacement and stresses are:

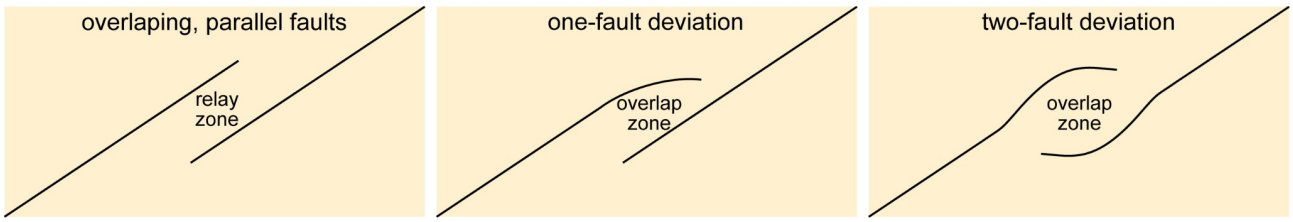
- mode 1 = opening mode: Tensile opening normal to the crack plane and propagation as extension fracture along the plane of the original crack;
- mode 2 = sliding mode = down-dip slip: Displacement parallel to the crack plane and normal to the crack tip-line with propagation as in-plane shear fracture along the plane of the original crack;
- mode 3 = tearing mode = strike-slip: Displacement parallel to both the crack plane and its tip-line with propagation as shear fractures.

Three basic modes of fracture



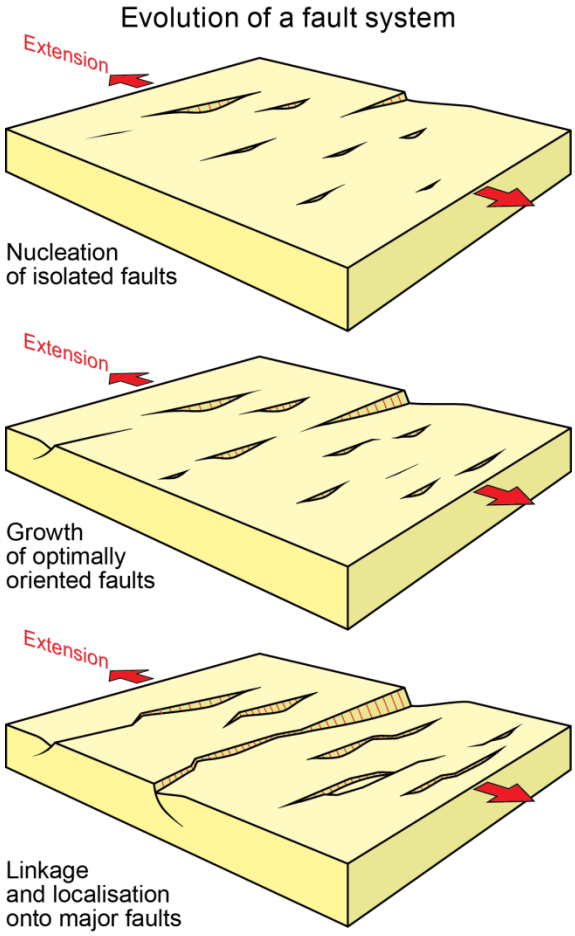
Propagation - Interaction - Coalescence

Fault systems evolve by the growth and **linkage** of smaller, individual fault segments. **Nucleation** describes the appearance of independent faults. **Propagation** describes the incremental growth of the fault dimensions and/or displacement. **Coalescence** describes the process by which initially isolated faults connect to form new and larger faults. Before faults are physically attached (**hard linkage**) to become structurally and mechanically coherent, they propagate towards one another. Their stress fields may already interact through an unfaulted relay zone of overlap (**soft linkage**) where elastic and ductile strain is concentrated.

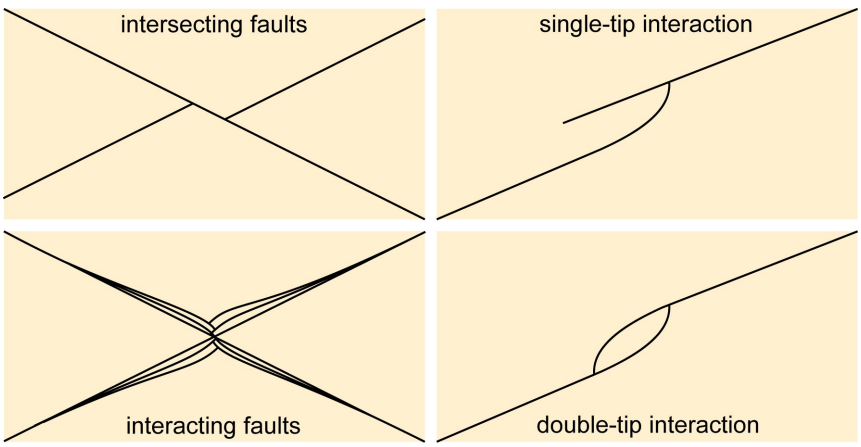


Main types of soft linkage

Hard linkage drastically changes the geometry of fault planes, which, conversely, do not significantly change during soft linkage. The fundamental concept is that growth and connection of individual fault segments take place as total displacement increases. The linkage and the resulting variations in displacement along main faults affect the location and thickness of syn-faulting sediments at various stages of fault development.



Stress anomalies, in particular at fault tips, locally perturb both the magnitude and direction of the regional stress field. The stress field in the overlap zone controls the growth of the overlapping faults, hence the linkage geometry. There are four main types of fault interaction: intersection, interaction, single-tip interaction, and double-tip interaction.



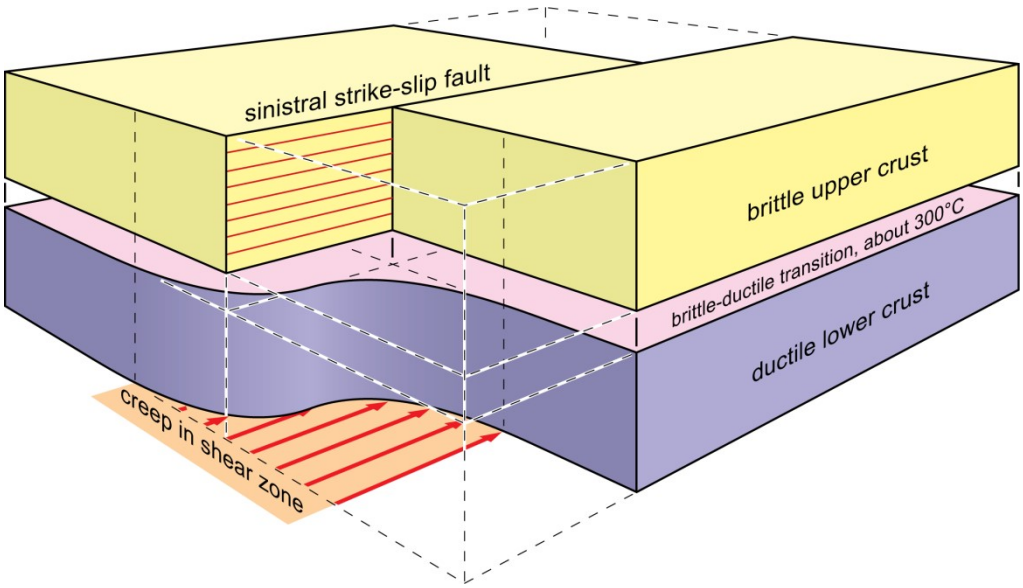
Main types of fault interactions

Displacement / Stress History

Rocks under stress undergo strain until they eventually break, forming a fault. Major faults show large total displacements that have accumulated incrementally over a long period. Their history encompasses interseismic periods of no-slip and seismic displacements.

Stable / unstable sliding

Fault creep is very slow slip, at typical average rates of a few cm/yr, without a perceptible earthquake. This steady, "aseismic" displacement probably takes place under near-constant shear stress within ductile shear zones and faults lubricated by clay minerals.



The observation that the slip of natural faults can occur with or without generating detectable seismic waves is paralleled by the laboratory observation that frictional sliding between rock surfaces can occur with or without detectable and sharp stress drops. **Stable sliding** occurs at a constant velocity without jerks and stress drops. **Unstable sliding**, also called **stick-slip behavior**, repetitively occurs with prominent jumps corresponding to episodic stress drops.

Stick-slip model

Seismic and interseismic faulting is compared to sliding of a weight resting on a flat and rough surface (plastic deformation) and pulled laterally through a helical spring (elastic deformation). The pull exerted by the spring must be larger than static friction to displace the weight.
The typical history is as follows (imagine also moving a piano):
If the surface is smooth, the weight glides at a more or less constant rate for a given stress, (an example is pushing something on ice).
If the surface is rough, motion consists of two stages:

- The applied force first slowly increases but nothing happens to the weight. Shear stress increases along the boundary between the weight and the slide surface as the system is loaded elastically, but static friction temporarily keeps the weight from moving.
- When the tension in the spring reaches a critical value, the weight leaps ahead. The stored elastic strain is released as the spring shortens and the basal shear stress drops. Back to memories of moving a piano, it is well experienced that it takes more effort to start shifting an object than to keep it moving once it is sliding. This is because the static friction sticking the weight to the surface is larger than the dynamic friction opposing motion when it starts.
- Once basal stress has dropped to below dynamic friction the weight is stuck again.
- If loading is maintained, the cycle of alternate jerky sliding and stress release starts over.

Coulomb law of sliding

Laboratory experiments on the sliding between two solids show that it follows, as a first approximation, the Coulomb law of sliding (not to be confused with the Coulomb fracture criterion). The mathematical expression is the same as equation (4):

$$\sigma_S = f(\sigma_N) \tag{4}$$

This is a macroscopic law. At the microscopic scale, the actual contacts only concern small fractions of the surface, with a total area proportional to σ_N. Microscopic processes will be different, depending on whether the bumps are overcome through elastic, brittle or ductile deformation. The coefficient f has a value at standstill (**static friction** coefficient μ_s) larger than during sliding (**kinetic coefficient** μ_k). Frictional sliding is such that:

(a) Either the shear stress equals or is lower than for static conditions, and there is no movement, or

(b) it is equal and there is sliding with an undetermined speed.

This slip condition (see equation 5) is linear and passes through the origin of a Mohr diagram. The Coulomb law of sliding accepts the relationship between shear stress and normal stress. Irregular interseismic times can be replicated by varying the ratio of the friction force and the normal force (the coefficient of static friction), which depends on some microscopic and often invisible properties (asperities, strength, stability, etc.) of the fault surface. Reducing the normal stress reduces the static friction to be overcome for sliding (unclamping effect), hence reduces the interseismic time. Conversely, increasing the normal fault clamps the model fault and increases interseismic time.

Frequency of events

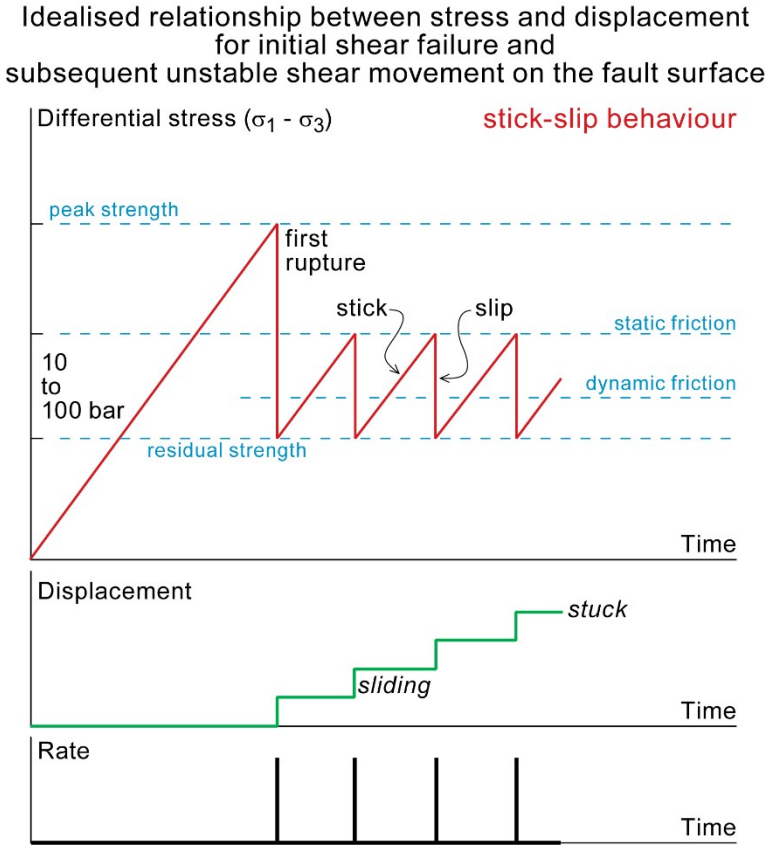
Stick-slip results from a familiar phenomenon: it is harder to start an object sliding than to keep it sliding. This is because the static friction against sliding exceeds the dynamic friction that opposes motion once sliding starts. The actual friction force that must be overcome depends on microscopic details of how rough is the gliding surface. This means that many parameters cause the weight to be stuck despite an applied force. All of these causes are metastable states because the stuck weight is in a stable position, yet not the lowest energy state since friction induces elastic strain in both the floor and the weight, and this strain corresponds to a certain amount of stored elastic energy. Among the metastable states, the set of configurations the weight will visit while performing its jerky motion has some particular importance. These states are **marginally stable**. A slight increase in the applied force can lead to almost any response: the same amount of driving force yields sometimes a small forward jump, sometimes a large one. The marginally stable states are believed to lack a typical time or length scale, which leads to a correlation function that describes the frequency with which events occur. This function is commonly a power law. The lack of a typical scale is similar to the configuration of a thermodynamic system at a critical temperature. This notion has led to the recent concept of Self-Organised Criticality.

> Reminder: The critical behavior of thermodynamic systems is well understood. For all temperatures, one can disturb the system locally and the effect of the perturbation will influence only the local neighborhood. If a transition temperature is reached, the local distortion will propagate throughout the entire system and something extraordinary happens (e.g. change of phase, solid to liquid for water at 0°C). The system is critical in the sense that all members of the system influence each other.

Stick-slip sliding on faults

It is now widely believed that the most important cause of earthquakes is large-scale stick-slip behavior on faults. The **seismic cycle** is subdivided into (1) preseismic, (2) coseismic and (3) postseismic phases:

1.) Stress and mainly elastic strain accumulate over a long preseismic period of tectonic loading until
2.) Frictional resistance along the fault is overcome when sudden and violent coseismic failure of the fault occurs until
3.) The stress drops to a value, the **residual strength**, at which friction prevents further slip. Aftershock activity decays inversely with time.
4.) Postseismic creep and the formation and collapse of fault-related dilatancy drives transient pulses of hydrothermal fluids.



As in the initial fracturing of rock specimens, there may be noticeable microfracturing activity associated with and premonitory of each of the stress-drop episodes during frictional sliding experiments. This possible 2.a) phase of preseismic anelastic deformation may involve foreshock activity and accelerating precursory slip.
Stress increase remains low and the upper limit of the saw-tooth stress/displacement relationship is a **threshold** known as the **sliding friction**. The usual magnitude of the stress drop associated with earthquakes is in the range of 10 to 100 bars, with a logarithmic mean at about 30 bars (3 MPa).

Faulting and fluid displacement

Fault 'valve' models have been proposed, where increasing fluid pressures below sealed reverse faults periodically trigger displacements with subsequently enhanced fluid flow and re-crystallization of the fault zone (silicification or calcitisation, etc). Fault 'pump' models suggest that the coseismic collapse of strain-induced micro-crack dilatancy within the host rock flushes fluid through the broken fault zone rock. Post-seismic creep and compaction of newly formed cataclasite, in turn, lead to pore volume reduction with fluid expulsion and vein emplacement in the adjacent wall rocks.

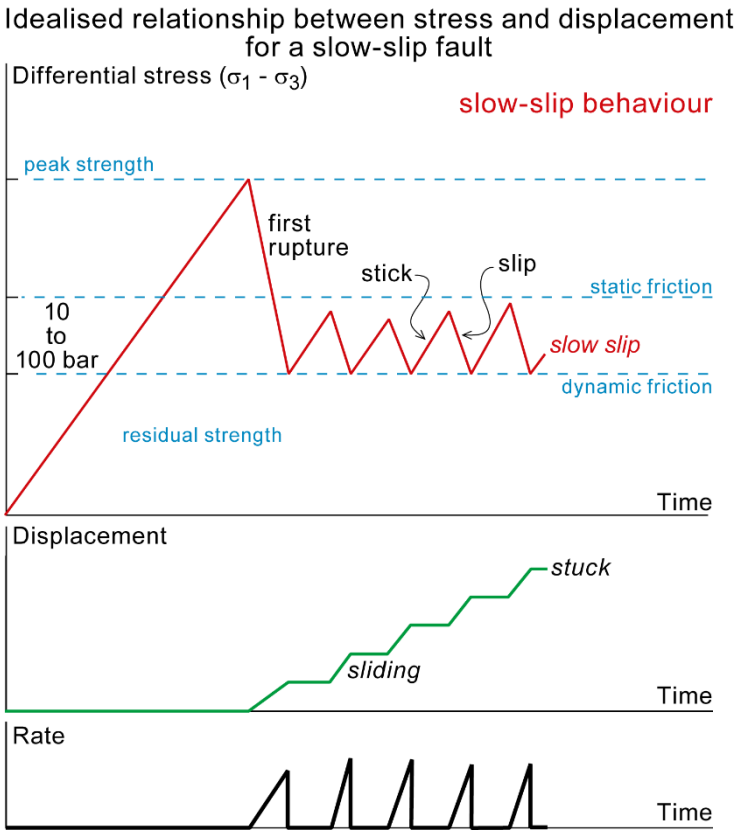
Variables influencing stick-slip behavior

Temperature, effective confining pressure, pore fluid chemistry, and rock type are among the variables that determine whether stick-slip or stable sliding occurs.

- High temperatures suppress stick-slip and this is one possible reason why earthquake foci are restricted to shallow depths (less than 20 km). Stick-slip is nearly absent at low temperatures under low effective confining pressure.
- At higher effective confining pressures, there is a positive effective normal stress on the fault plane and stick-slip behavior appears, the magnitude of which is dictated by the frictional resistance to sliding on the fault plane. The effects of pore fluid chemistry, as opposed to pore fluid pressure, are not yet clear, but important effects may exist if stress corrosion at crack tips is a factor governing the growth rate of microfractures.

Slow slip: slow earthquakes

Instrumental records, often combining a range of methods (seismic, GPS, long-term strain measurements in boreholes), indicate that other processes than dynamic rupture (i.e. seismic rupture lasting seconds to minutes) release the strain energy accumulated in the Earth's crust. "Aseismic" slip (i.e., rupture propagating over hours to several months) describes several types of usually discontinuous events such as episodic, non-volcanic **tremor** or almost imperceptible, low-frequency rumblings.

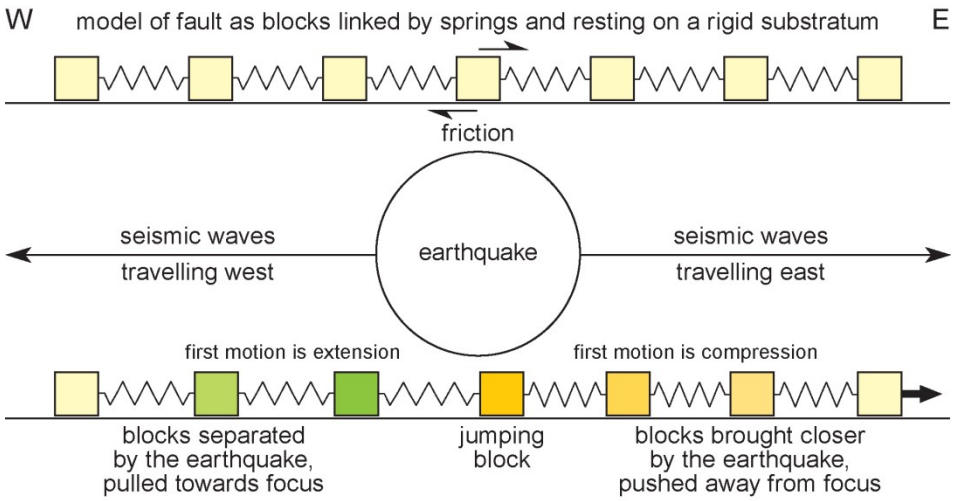


Slow slip is still poorly understood and its occurrence vs. faster seismic rupture is variably attributed to the fault structure (e.g. shape and asperities), fault material (e.g. gouge composition, fluid pressure) and depth (rather shallow in the crust). Slow slip involves deformation rates expected to develop brittle-ductile structures in the fault zones, which geologically remain a discontinuity in the crust. The process has been documented for all types of faults and tectonic regimes.

Fault segments

Friction forces vary at different points on a fault. The instantaneous rates of slip may, therefore, strongly fluctuate, ranging from zero to rates comparable to the elastic wave velocity of the surrounding rock (ca 7 km.s^{-1}). Thus, an earthquake may emanate from the fault portion where the slip rate is temporarily very high while, simultaneously, another portion of the same fault is locked or is

only exhibiting slow, aseismic creep. An analogy is provided by a series of blocks connected by springs and pulled from one extremity. Each spring represents a segment. The complex interplay of relative movements between blocks simulates the diachronous and segmented movement on the fault.



First motion of a one-dimensional earthquake
(modified after Cox & Hart 1986 *Plate tectonics. How it works*. Blackwell Scientific Publications)

The average slip rate of major faults over long periods is of special interest in tectonics. It may be estimated geologically by using several methods to obtain the displacement and dividing by some estimate of the elapsed time. Average slip rates may also be obtained on some faults (e.g., oceanic transform faults) by applying the theory of sea-floor spreading. Two methods are available for making relatively short-term determinations of slip rate. The first is to carry out repeated surveys of networks of geodetic stations either side of a fault. The second is a simple seismological method in which the surface wave magnitudes of all earthquakes occurring on a given fault in a given interval of time are used to estimate the total slip on the fault during that time and, hence, the average slip rate. Calculated slip rates are in approximate agreement with geodetic measurements. In some cases, the slip rates indicated by this method are lower than those indicated geodetically, suggesting either that some additional slip is occurring by creep or that strain is accumulating for an earthquake. Besides, structures and topographic features observed along faults indicate that the pre-faulting strain is not entirely recovered during a seismic event. Aftershocks also indicate that the total amount of released strain/energy is not instantaneous.

Sliding along existing faults

In natural conditions, failure is likely to take the form of sliding along pre-existing fractures rather than opening new ones. Sliding along pre-existing fractures does not have to overcome cohesion, only sliding friction. In such conditions, the relationship between σ_S and σ_N is equation (5):

$$\sigma_S = \sigma_N \tan\phi \qquad (5)$$

a line that goes through the origin in a Mohr diagram, and, therefore may intersect a Mohr circle of a given state of stress in two points. Sliding will be possible along any surface oriented at angles with σ_3 that corresponds to the intercepted arc of the Mohr circle. This relationship shows that increased pore pressure, which shifts Mohr circles to the left, may reactivate existing fractures.

Exercise

Draw a Mohr circle representation with σ_1 = 100 MPa, σ_3 = 50 MPa, for a rock with a cohesion strength of 20 bar and a failure surface whose slope is 25°. Draw the same body with a pre-existing fracture.

As the coefficient of friction is relatively well defined for most rocks, it can be shown that frictional sliding will occur on pre-existing faults when $\sigma_1/\sigma_3 \approx 3$. Faulting is controlled by the vertical principal stress σ_v, and one horizontal principal stress, σ_3 in extension and σ_1 in compression. In cases of hydrostatic pore pressure, these relationships show that in extensional areas $\sigma_3 \approx 0.6\,\sigma_v$, in compression areas $\sigma_1 \approx 2.3\,\sigma_v$, and in strike-slip faulting areas, when $\sigma_v \approx (\sigma_1 + \sigma_3)/2$, then $\sigma_1 \approx 2.2\,\sigma_3$. In situ stress measurements to depths of about 2 km at intraplate sites have confirmed these simple equations.

Cataclastic flow

Fluid overpressure enhances hydraulic fracturing and **cataclastic flow**. **Cataclasis** refers to grain fracturing while friction between grains is sufficiently large to inhibit sliding on grain boundaries. Cataclastic flow refers to deformation during which distributed microfractures at grain scale produce clasts that frictionally slide past each other and rotate. Grain crushing happens at such a minute scale that the macroscopic structure of the rock seems to results from flow in a ductile manner. However, microfractures and grain rotation are brittle, cataclastic features. Consequently, fault movement takes place across a commonly narrow, cataclastic fault zone. Owing to these numerous microcracks, hence porosity and permeability, fault zones tend to channelize the fluid pathways. Channelized fluid flow weakens the faults considerably, which accounts for the concentration of brittle deformation into highly localized planar discontinuities.

Locking

Many cataclastic processes may increase the number of **asperities**, i.e. points and areas where fault rocks are in contact. The total area of contact may cause high resistance to sliding, to the level where **adhesion** prevents further faulting. Adhesion is mechanical, due to interpenetration and locking of fault blocks, and chemical due to secondary crystallization bonding surfaces along the fault zone.

Faulting and earthquakes

Studying fault movements and faulting is important to get a better understanding of earthquakes because earthquakes are thought to release strain and stress on pre-seismically locked faults.

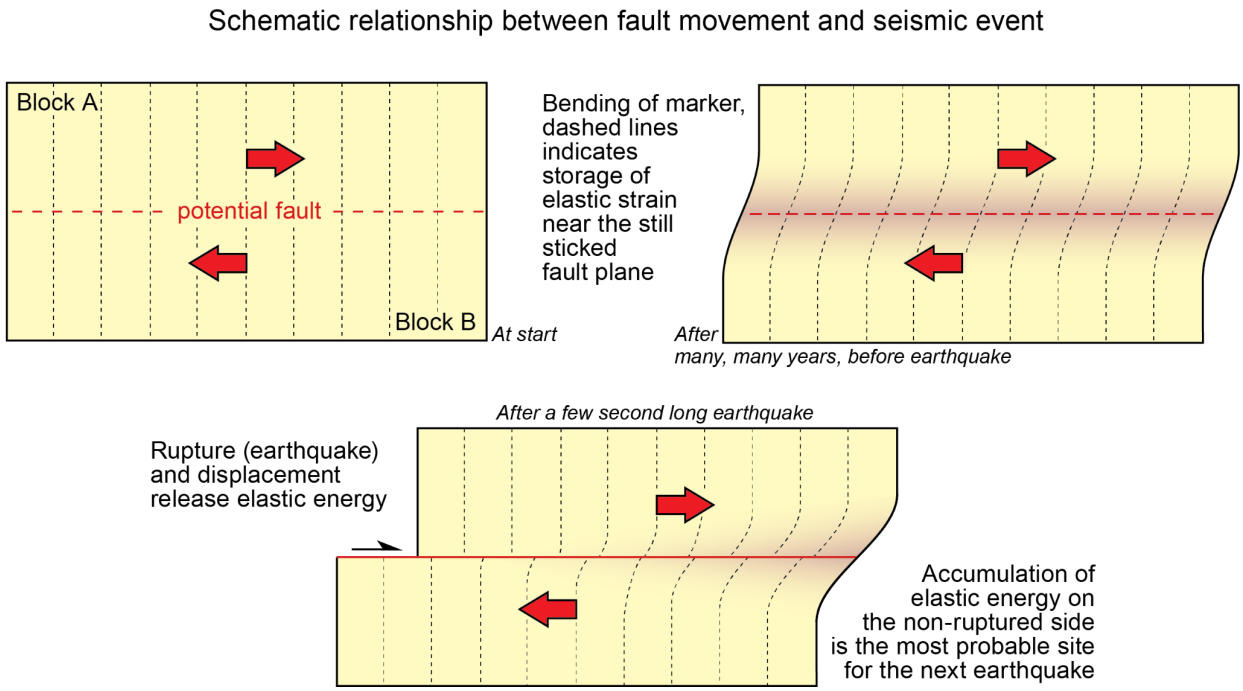
Elasto-plastic model

Earthquake reconstruction uses a spring pulling a mass resting on a rough surface. The pull on the spring must exceed the frictional resistance to sliding of the mass for the latter to move. Before any slip, one can pull the spring back and forth; this is elastic behavior. But once the mass slips it does not move back (the displacement is permanent) while the spring recovers some or all of its strain.

Model relevance

Plate movements entail relative movement between adjacent crustal blocks. Consequently, two reference points taken away, on both sides of the bounding fault zone move with respect to each other. These relative movements are continuous in the lower, ductile crust where fault creep is permanently active. The lower crust carries the upper, brittle crust in which the shallower section of the fault does not instantaneously follow the deeper, steady, ductile movement. The brittle fault section is locked while elastic strain accumulates with a gradient centered on the fault plane (a straight line across the fault, initially joining the two distant points, becomes sigmoidal). This portion of the seismic cycle, during which strain and stresses increase with time along the fault, corresponds to the **interseismic interval**. When the local stresses exceed the frictional strength of the fault zone, slip suddenly occurs with an earthquake along a newly created or reactivated fracture. The upper crustal rocks instantaneously release their accumulated elastic strain (the sigmoidal line snaps back to its original straight shape but is broken by the fault). This **coseismic** strain reduction is the **elastic rebound**. A

new **earthquake cycle** begins and will end with the next seismic slip. The rapid coseismic slip increases stress in the ductile crust. As these stresses are relieved after the earthquake, either through aseismic fault slip or through more broadly distributed viscous flow at depth, the upper-crustal fault is re-stressed and the fault region re-strained. This phase corresponds to the **postseismic** period, which grades into the steady interseismic stress accumulation progressing to the next earthquake.



Accordingly, faults are seismic sources. The elastic behavior of rocks and the frictional properties of faults govern brittle deformation. The earthquake cycle provides an important framework for forecasting earthquakes and thereby mitigating their effects. In theory, one can predict the next earthquake on a particular fault from the deformation that accompanied the previous earthquake. At hand are the slow rate of permanent deformation determined from geologic observations (millimeters to centimeters per year), the repeat times between earthquakes on the same fault segment (earthquake recurrence, ranging from tens to several thousands of years) and the rate of interseismic strain accumulation determined from geodetic observations. In practice, this prediction has proven quite difficult because of the generally long time between earthquakes and the relatively short history of geodetic measurements. Besides, the cyclic deformation is highly idealized. Because of the paucity of appropriate observations, it is still not known whether interseismic strain accumulation or permanent deformation can be characterized by uniform rates on the time scale of individual earthquakes. In addition, fault properties and irregularities change during each seismic event, so that the total friction resistance to movement changes and alters the interseismic interval. In any case, the identification of seismic sources in an area is the first step in evaluating the earthquake risk. Because rocks have typical strength and plate tectonic frameworks have rather steady rates, particular faults tend to generate **characteristic earthquakes** with the same maximum magnitude.

First motion

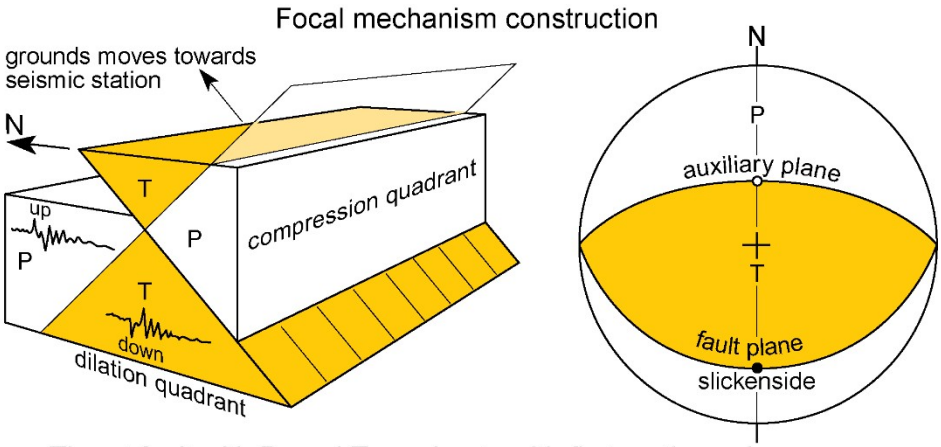
Seismic waves are vibrations generated by the rupture and sudden movement of rock at the **hypocenter**, or **focus**, of the earthquake. Two types of elastic, seismic waves are produced:

- Body waves radiate through the Earth's interior in all spatial directions from the source region.
- Surface waves propagate along the Earth's surface.

P (primary) and S (secondary) body waves have different characteristics in different directions and travel within the Earth faster than surface waves. Therefore, P waves are the first to arrive from a distant earthquake. The first motion they impel to the ground (the **impetus**) is divided into two classes:

- Upward, thus away from the earthquake source, which means that the related fault movement has pressed and pushed away all material points along the expansion direction from focus to the record point.
- Downward, thus towards the earthquake source, which means that the related fault movement has depressed and pulled in all material points in the contraction direction of propagation.

Hence, the direction of the first ground motion at a particular place depends on its position relative to the hypocenter and the type of faulting that occurred, i.e. whether the fault is moving rocks toward or away from the direction of the seismic station. In conclusion, the **first-motion** study of earthquakes can be interpreted in terms of fault orientation and displacement direction.



Thrust-fault with P- and T-quadrants with first motion seismograms and associated, lower hemisphere stereographic projection
first motion up = push away from epicenter; first motion down = pull toward epicenter

S waves produce a sideways shearing motion at right angles to the direction of propagation. Surface waves are the most damaging and consist of a complex horizontal and rolling (Rayleigh) motions.

Focal mechanism solutions

The direction of elastic forces released during an earthquake can be determined from the first motions recorded at as many seismograph stations as possible around the earthquake.

Concept

Upward first motion releases compression at the recording point; conversely, downward first motion expresses extension (dilation). Neither push (compression) nor pull (dilation) occurs along the fault plane because slip is shearing motion only. This is also valid for the non-material **auxiliary plane**, which is normal to both the fault plane and the slip direction. The fault plane and the auxiliary plane (the **nodal planes**) divide the space around the focus into four alternate **quadrants** where the first motion is a push or a pull. This **quadrant distribution** is characteristic of most natural earthquakes.

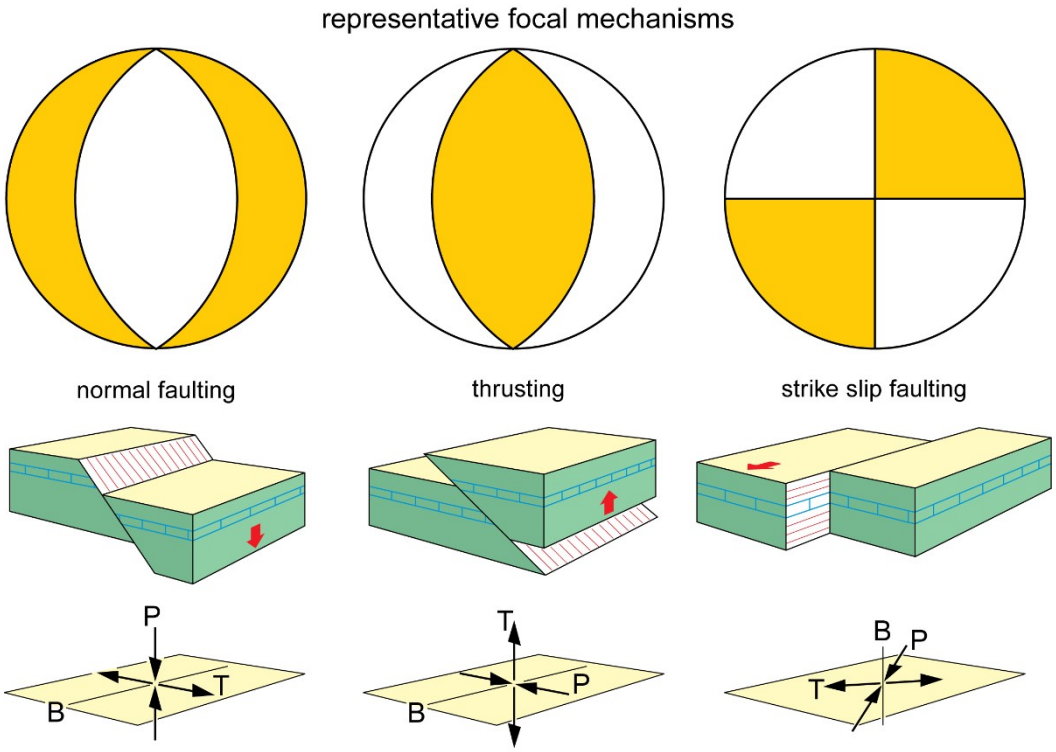
Implementation

The goal is to have the first compressional/dilational ground motion recorded by a sufficient number of widely spaced seismograms in different directions from the earthquake. Seismic stations recording very small or no first motion align along **nodal lines**. Generalized to three dimensions, these lines construct the nodal planes.

The standard procedure is as follows:

- Plot on lower-hemisphere, equal-area projections the oriented station points (considering wave trajectories in the spherical Earth) where the first P waves had push (downward first-motion of P traces on a seismogram) or pull (upward first-motion of P traces on a seismogram) character.

- Identify the pair of orthogonal nodal planes that best divide the space around the focus into compressive and dilatative quadrants.
- If faulting caused an earthquake, then one of the nodal planes is the fault plane and the other is the auxiliary plane.



Orientation of the principal stresses
P (maximum, compression) T(minimum, extension) and B (intermediate)

The resulting plot is a **focal-mechanism (**or **fault plane) solution**. The direction of displacement lies within the fault plane and is perpendicular to the auxiliary plane. It is the pole to the auxiliary plane. Hence, the slip direction and type of movement along the fault can be determined easily from the arrangement of compressional (conventionally shaded) and dilatational (shown blank) quadrants. Note also that the radius of the focal sphere has no value.

Application

The direction of compression (P-axis, the bisector of dilational quadrants) and the direction of tension (T-axis, the bisector of compressional quadrants) derived from the quadrant distribution are information on the stress field responsible for the considered earthquake. Accepting that P and T axes approximate the directions of maximum and minimum principal stresses, respectively, one can define the nature of the fault.

One problem is that a solution for an earthquake yields two nodal planes that both are possible faults since both have the same first motion. This ambiguity is an inherent characteristic of focal mechanism solutions based solely on P-wave impetus. In principle, it can be overcome by the study of S-wave characteristics or by other means. In practice, geological knowledge and/or observation, especially the orientation of surface breaks associated with an earthquake, quickly help to distinguish the fault plane from the auxiliary plane.

Another problem is that a pre-existing fault plane controls the seismic radiation pattern more than the in situ stress field: the earthquake focal plane mechanism always has the P- and T-axes at 45° to the fault plane and the *B* axis in the plane of the fault.

To use earthquake focal mechanisms to determine stress orientations, one considers average P-, B-, and T-axes for earthquakes occurring on different faults within a limited region.

Furthermore, the amplitude and frequency content of seismic waves only give information about the magnitude of stress released in an earthquake (**stress drop**) and not the absolute stress levels. In general, stress drops are about 1–10 MPa, a very small fraction of the shear stress that causes the earthquake.
In conclusion, a fault-plane solution is a method for using the seismograms from an earthquake to study the geometry and sense of motion on faults. This method permits interpretations of present-day movements on deeply buried or otherwise concealed faults, especially in the oceans, and has proved important by enabling the relative motions of lithospheric plates to be determined.

Attention: News media generally report the location of the **epicenter**, which is the point on the surface of the Earth directly above the focus; geophysical information includes the location of the epicenter and the depth to the focus. A focal mechanism solution is obtained by first determining the location of the focus.

Magnitude and fault movement

Geophysical information also includes the **magnitude**, a measure of the strength of the earthquake. The **Richter magnitude** (M) of an earthquake was first determined as the largest amplitude A of seismic waves recorded on a seismogram normalized to a local reference amplitude A_0 all expressed on a logarithmic scale and multiplied by an empirical factor of 3:

$$M = \log\left(A/A_0\right)$$

Since the Richter magnitude is a logarithmic scale, a whole number unit increase on the scale represents a ten-fold increase in vibration amplitude.
The **Gutenberg-Richter power law** states that the average number n of earthquakes with magnitudes ≥ M recorded per year by a particular seismic network is a decreasing exponential function of M:

$$\log n = a - bM$$

with a and b constants characteristic of the particular region. This law describes the frequency of different sized earthquakes. In most regions, b is about 1 so that earthquakes generally become nearly 10 times more frequent for every unit decrease in magnitude (for instance, there are 10 times fewer earthquakes of magnitude 5 than earthquakes of magnitude 4). The Gutenberg-Richter rule is acceptably valid for earthquakes at epicentral distance < 600 km and magnitude M < 7.
Magnitude is a measure of the amount of energy released, which is proportional to the fault displacement. For earthquakes of magnitude >7, the **moment magnitude** (M_W) is used. It accounts for the average amount of slip on the fault that produced the earthquake, the rupture area and the shear modulus of rocks that failed. The Richter and moment magnitude scales are comparable over the range of magnitude 3 to 7.
The energy E released during an earthquake is given in Joules by:

$$E = 10^{11.8+1.5M}$$

A magnitude increase of 2 is equivalent to an energy increase of approximately 1000 times.

Displacement and rupture length

The ratio between seismic slip u_s and the rupture length L_s is considered to be:

$$10^{-4} > u_s/L_s > 10^{-5}$$

The ratio between total displacement u_t and total fault length L_t is estimated as:

$$u_t/L_t \approx 0.1$$

Case studies have shown that the rupture point propagates along the fault plane, generally and mostly in one direction, unzipping locked portions of the fault at a rate of ca 3km/s.

Earthquake intensity - Material amplification

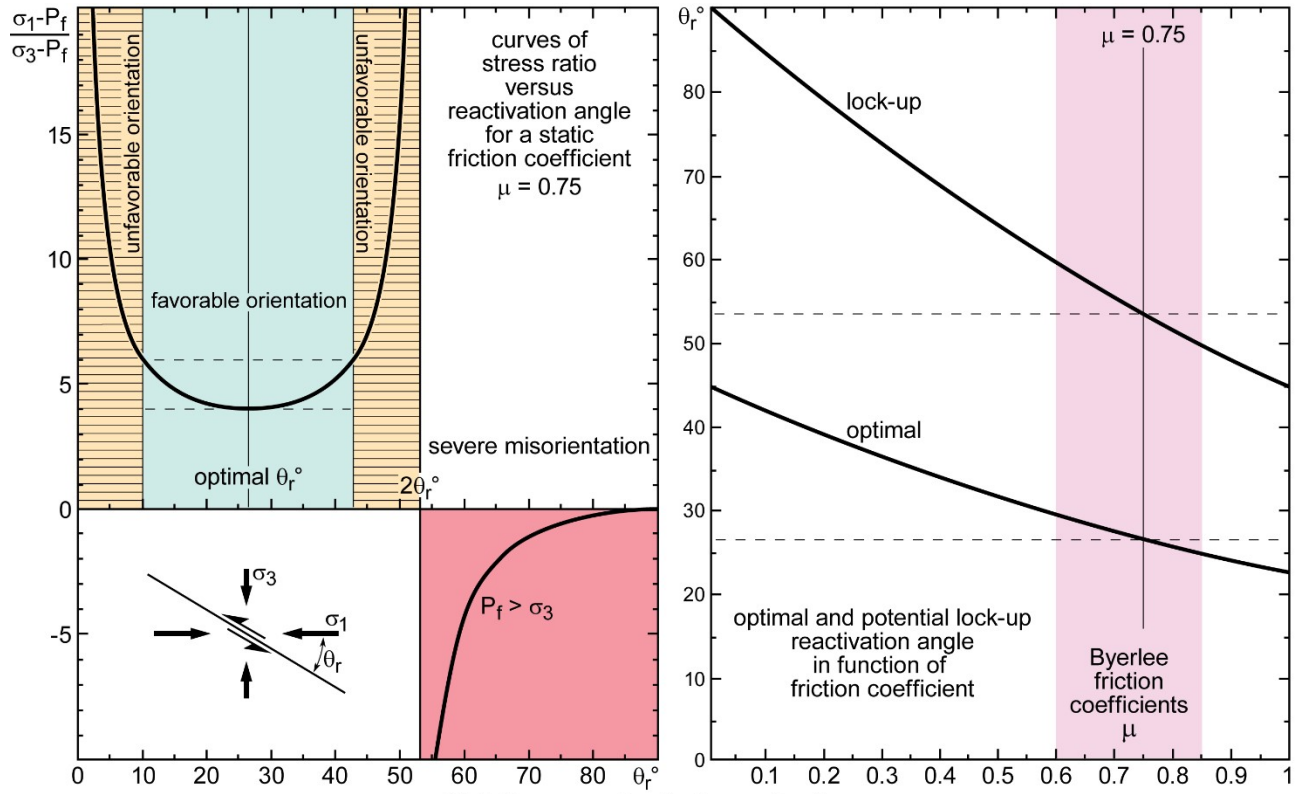
The **intensity** of an earthquake is related to local damage, duration and other effects on structures and people. Earthquake intensity is scaled from I to XII on the qualitative **Modified Mercalli scale**. Intensity varies with distance from the epicenter and from place to place within the shaken area.
Earth materials have different elastic properties and, therefore, respond differently to seismic waves. The amplitude of shaking is increased in unconsolidated sediments because the energy of a seismic wave remains practically constant when passing from one rock type to another. Since the seismic velocity is lower in unconsolidated than in consolidated rock, keeping energy constant implies that the displacement amplitude of the waves must be bigger. This **material amplification** is important for vulnerability evaluation. It is often the reason why buildings on unconsolidated rock are destroyed while buildings in the surrounding consolidated rock are not. Besides, shaking may produce an increase in pore-water pressure of water-saturated weak sediments, which results in sediment liquefaction and flow.

Slip rate and recurrence

Any fault has a long, discontinuous movement history. The slip rate is the ratio of the amount of displacement to the time interval over which that displacement took place. The average recurrence interval is the average time interval between earthquakes.
However, these may vary in time, which casts suspicion on average values. Frequently, earthquakes are clustered events separated by relatively long periods of quiescence. The duration of the period between two earthquakes along the same fault has been attributed to several physical parameters such as the relative velocity between the two walls of the fault, the mineralogy, and anisotropy of the fault walls and the fault rocks, the morphology of the fault plane, the thermal state, the fluid pore-pressure, and others.

Reactivation

Recorded seismicity is largely due to the reactivation of existing fault planes. Conditions for reactivation depend on frictional properties, which in turn depend on fluid pressure and the fault orientation with respect to principal stresses. Two-dimensional analyses with σ_2 parallel to the strike of the fault show that reactivation is relatively easy for an effective stress ratio $(\sigma_1\text{-}P_f)/(\sigma_3\text{-}P_f) \approx 4$ and low angles (ca 15 to 45°) between the fault plane and σ_1. Faults at a high angle are locked unless the fluid pressure is sufficiently high to virtually eliminate friction. It is easier to create new slip planes if existing faults have unfavorable conditions.



Exercise

Combine equations (3) and (5) to express the stress ratio in function of the angle between the fault plane and σ_1; draw the curves for friction coefficient 0.6 and 0.85.

Conclusion

Experimental results can be summarised as follows:

- Fractures are generated when the effective stress exceeds the tensile strength of the rock.
- At a given temperature and strain rate, increased confining pressure increases both the yield stress and ultimate strength;
- At a given confining pressure, increased temperature or decreased strain rate lowers both yield stress and ultimate strength.

In short, brittle deformation is highly pressure-sensitive.

The simple and widely applicable Coulomb criterion predicts a linear strength-pressure relationship. At higher confining pressures where ductile behavior begins, or at very low values of σ_3 where other special effects enter, plots of σ_N against σ_S at failure are characteristically concave toward the σ_N axis. It is the empirical Mohr failure envelope tangent to Mohr circles at different σ_S and σ_N at failure. This line separates stable and unstable stress states, the latter giving rise to brittle deformation.

The mechanistic Griffith failure criterion describes the tensile failure. It assumes that real materials contain slit-like imperfections (Griffith cracks) that open, propagate and link, which leads to macroscopic rupture.

The increase in fluid pressure reduces the effective stresses and brings the stress state closer to the failure envelope.

Faulting occurs to relieve accumulated shear stress on faults. The sudden movement of strained crustal blocks that break along new or pre-existing faults causes earthquakes. Earthquake displacement does not occur over the whole length of a large fault: a small area of a fault displaces at one time with other

areas displacing at another time. However, earthquake displacement at one particular area along a fault may occur at regular recurrence intervals.
The deformation of the upper crust is mostly accommodated by important displacements along fault zones. Such large displacements result from seismic slip increments representing reactivation events of faults over lengthy periods. Stress cycling accompanying episodic reactivation is likely to be the norm in the vicinity of major faults whose seismic cycles involve long periods of elastic strain and stress accumulation, driven by aseismic ductile deformation at depth, ultimately released by sudden fault slip events. As a corollary, earthquake focal mechanisms are the most ubiquitous indicators of stress in the lithosphere.

Recommended literature

Hubbert M.K. & Rubey W.W. - 1959. Role of fluid pressure in mechanics of overthrust faulting: I. Mechanics of fluid-filled porous solids and its application to overthrust faulting. *Geological Society of America Bulletin.* **70 (2)**, 115-166, 10.1130/0016-7606(1959)70[115:ROFPIM]2.0.CO;2

Jaeger J.C. - 1969. *Elasticity, fracture and flow: with engineering and geological applications*. third, Methuen & Co LTD and Science Paperback, London. 268 p.

Jaeger J.C. & Cook N.G.W. - 1979. *Fundamentals of rock mechanics*. 3, Chapman and Hall, London. 593 p.

Jaeger J.C., Cook N.G.W. & Zimmerman R.W. - 2007. *Fundamentals of rock mechanics. Fourth edition.* Blackwell Publishing, Oxford. 475 p.

Jolly R.J.H. & Sanderson D.J. - 1997. A Mohr circle construction for the opening of a pre-existing fracture. *Journal of Structural Geology.* **19 (6)**, 887-892, 10.1016/S0191-8141(97)00014-X

Mandl G. - 1988. *Mechanics of tectonic faulting*. Elsevier, Amsterdam. 407 p.

Mandl G. - 1999. *Faulting in brittle rocks*. Springer, Berlin. 434 p.

Savage J.C., Byerlee J.D. & Lockner D.A. - 1996. Is internal friction friction? *Geophysical Research Letters.* **23 (5)**, 487-490, 10.1029/96GL00241

Scholz C.H. - 1992. *The mechanics of earthquakes and faulting*. 2nd, Cambridge University Press, Cambridge. 439 p.

Sibson R.H. - 1989. Earthquake faulting as a structural process. *Journal of Structural Geology.* **11 (1-2)**, 1-14, 10.1016/0191-8141(89)90032-1