GEOPHYSICAL RESEARCH LETTERS, VOL. 17, NO. 12, PAGES 2109-2112, NOVEMBER 1990

FRICTION, OVERPRESSURE AND FAULT NORMAL COMPRESSION

J. Byerlee

U.S. Geological Survey, Menlo Park, CA 94025

Abstract. More than twenty–five years ago Miller and Low reported the existence of a threshold pore pressure gradient below which water would not flow through clay. Recent experimental observations of the shear strength of structured water on biotite surfaces have provided a physical basis for understanding this threshold gradient. The existence of this phenomenon has profound implications for the rheological properties of mature fault zones, such as the San Andreas, that contain large thicknesses of fault gouge. For example, a clay–filled fault zone about 1 km wide at the base of the seismogenic zone decreasing to zero width at the surface could support core fluid pressure equal to the maximum principal stress over the entire seismogenic zone. As a result, the fault would have near–zero strength and the maximum principal stress measured on the flanks of the fault, would be oriented normal to the fault surface. Another consequence of the threshold gradient is that normal hydrostatic fluid pressures outside the fault zone could coexist with near–lithostatic fluid pressures in the interior of the fault zone without the need for continual replenishment of the overpressured fluid. In addition, the pore pressure at any point should never exceed the local minimum principal stress so that hydrofracture will not occur.

Introduction

Recent determinatons of the stress directions indicate that the maximum principal stress σ_1 is almost normal to the San Andreas fault [*Mount and Suppe*, 1987; *Zoback, et al.*, 1987]. This is consistent with the low shear strength of the fault inferred from the heat flow data. These observations have proven to be very difficult to explain in terms of our conventional understanding of fault rheology. In most regions of California, faults exist with almost all orientations. If all the faults in a region have approximately the same coefficient of friction μ, then elementary coulomb analysis demands that the maximum angle that σ_1 can make with the most active fault is 45°, even if μ is zero. It has been suggested, [*Zoback et al.*, 1987] that μ for sliding on the San Andreas may be very low, but high for the subsidiary faults. Calculations (*A.H. Lachenbruch*, personal commun., 1990) show that μ on the San Andreas needs to be ≤ 0.1 to satisfy the heat flow constraints. The weakest mineral that can be expected to occur in any significant quantity in fault zones is montmorillonite clay. In a series of experiments to measure the strength of some water saturated clays [*Radney and Byerlee*, 1988], we found that the effective stress law for frictional sliding was obeyed and that the μ for montmorillonite was $\geq$0.2. Thus even if the San Andreas fault was filled with pure montmorillonite, which seems unlikely, the heat flow constraints would be exceeded.

Paper number 90GL02295
0094-8276/90/90GL-02295$03.00

The only way for the fault to have the necessary low strength is for pore pressure to exceed 75 percent of the lithostatic stress in the San Andreas fault zone while pore pressure follows the hydrostatic gradient everywhere else (*A.H. Lachenbruch*, personal commun., 1990).

We will denote an effective stress component $\bar{\sigma}_i = \sigma_i - p$ where p is fluid pressure. Then, for fluid pressures following the hydrostat it is commonly found [*Zoback and Healy*, 1989] that the maximum and minimum effective principal streses are related by $\bar{\sigma}_3 = \bar{\sigma}_1/3$. If σ_1 is approximately equal to the lithostatic stress and the density of the overlying rock is 2,500 kg/m^3, then $\sigma_3 = 0.6\,\sigma_1$. Thus, from conventional rock mechanics we would expect hydrofracture to occur if λ, the ratio of pore pressure to lithostatic stress, in the fault zone exceeds 0.6. This was also pointed out by *Zoback et al.* [1987].

If the country rock is so strong that hydrofracture does not occur, then it is necessary for the country rock to have very low permeabilty to prevent the high pressures in the fault zone from bleeding off. Calculations by *Hanshaw and Bredehoeft*, [1963] show that even if a high pore pressure region is surrounded by material 1 km thick, with a permeability of $10^{-21} m^2$ (which is an exceedingly low permeability for geological materials) λ can not be maintained above 0.75 for more than 10,000 years. This is because the specific storage, which is a function of the pore volume and the compressibility of water and rock, is so low that only a very small amount of water needs escape to drastically lower the pore pressure.

A clue to explaining the apparent paradox of the mechanics of large fault systems comes from crustal drilling projects and the oil exploration industry. It has been found that overpressures are common in the earth even in rocks as old as Cambrian. In fact, in some cases the pore pressure can even exceed the lithostatic stress [*Fertl*, 1976]. In some oil fields these very high pressures in oil or gas can be explained by the very high capillary pressures that are required in the non–wetting phase to displace water from the narrow channelways in the overlying rocks [*Watts*, 1987]. In many places, however, water is the only liquid phase present, so we cannot appeal to capillary forces alone to explain the overpressure.

Threshold Gradient

An explanation of how overpressure can be maintained for long times comes from the results of laboratory experiments carried out in the field of soil science. It has been found that at very low pressure gradients water will not flow through dense clay. This is shown schematically in Figure 1. For Darcian flow the flow rate increases linearly with pressure gradient. The slope of the straight line through the origin is a measure of the permeability of the material. With dense clay the flow is non–Darcian at low pressure gradients. In this case

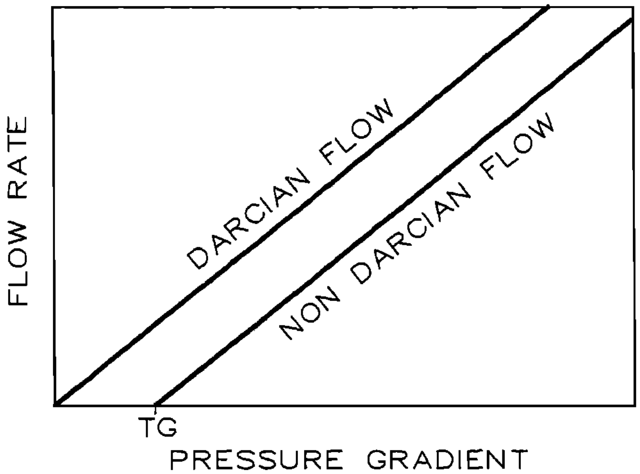

Fig. 1. A schematic diagram of flow rate as a function of pressure gradient for flow through a Darcian material, and a non darcian material with a threshold gradient TG.

the zero intercept on the pressure gradient axis is called the threshold gradient (TG). At pressure gradients below the TG, flow does not occur so that the permeability of the material is zero. The experiments to measure the TG are very difficult to carry out, and some hydrologists have assumed that the observed TG is an experimental artifact mainly because in some experiments no TG is observed, whereas in others it is, even with the same material although of different porosity. In addition, until now there has been no satisfactory physical explanation for the phenomena. It has been suggested *Lutz and Kemper* [1959], that the streaming potential generated by the flow can produce small counter currents along the pore walls in a direction opposite to the main flow, and it may be possible for this mechanism to reduce the flow rate at low pressure gradients, but it can not explain why flow ceases altogether below the TG. Others [*Miller and Low*, 1963, *Von Engelhardt and Tun*, 1955 and *Swartzendruber*, 1962] have suggested that non–Darcian behavior can be attributed to a non–Newtonian shear rate dependent liquid viscosity caused by clay–water interactions. But, here again, this mechanism can not explain why flow ceases althogether below the TG. *Hansbo* [1960], *Mitchell and Younger* [1967], have suggested that there may be a particle rearrangement at high pressure gradients to allow water to flow only when the TG is exceeded, but it is difficult to see how this mechanism would operate in dense clay.

A sound physical explanation for the TG phenomena comes from the results of recent experiments carried out in the field of surface science. It has been found that when two surfaces, separated by a liquid are brought closer together than about 10 times the diameter of the molecules, the liquid becomes ordered into discrete layers. When this occurs the liquid has a finite strength. Furthermore the normal stress required to expel a layer of the liquid increases as the number of layers decreases. If a shear stress is applied, shear displacement will not occur between the surfaces until the shear stress reaches a threshold level. The magnitude of the threshold shear stress increases as the number of layers of the liquid decreases. In addition, the shear stress required to shear the liquid is independent of the shear velocity, [*Israelachvili et al.*, 1988]. With thin films of a liquid between atomically smooth surfaces the concept of Newtonian viscosity breaks down, and attempts to calculate an effective viscosity lead to values that are many orders of magnitude larger than the bulk values. This applies to water as well as all the other liquids so far studied, [*McGuiggan et al.*, 1989]. In the case of water its shear strength is directly proportional to the normal stress that can be supported [*Homola et al.*, 1989].

It should make no difference whether shearing of the water is caused by a shear displacement of one of the surfaces relative to the other or whether the two surfaces are held fixed and the water is forced to move between them by a pressure gradient. The first case requires a threshold shear stress and the second case, requires a threshold gradient. Because the threshold stress is zero for greater than 10 layers of molecules it would be expected that the threshold gradient would not occur when the water–clay ratio is large, but would be manifest when this ratio is low, as found by *Miller and Low*, [1963]. With dense clay it has been found that the TG starts to be measurable when the void ratio, the volume of the voids to the volume of the solids, is about 2 and increases to 0.5 MPa/m at a void ratio of about 0.75 [*Ping*, 1963].

From physical considerations all that should be required for a TG to exist in fault gouge is that the width of the passageways occupied by water not exceed ten times the diameter of a water molecule. If this is true, then it may not be necessary for the walls of the passageways through which the water moves to be perfectly parallel as they are in clays, such as montmorillonite, illite, serpentine, etc. Thus a TG may also exist in a gouge with particles of clay size regardless of their composition, although experiments are required to confirm this. In any case it does appear reasonable at this stage to assume that in many gouges the TG is real and we can proceed to investigate whether the existence of TG leads to a resolution of some of the hitherto puzzling tectonic observations.

Model

In our model of the fault zone (Figure 2) we assume that in country rock the effective maximum principal stress, $\bar{\sigma}_1 = \sigma_1 - p_h$ is equal to three times the effective minimum principal stress $\bar{\sigma}_3$ where p_h is the hydrostatic pore pressure. σ_1 is assumed to be normal to the plane of the fault containing gouge $2D + d$ thick. For equilibrium σ_1 must remain constant through the gouge zone. The TG is assumed to be $(\sigma_1 - p_h)/D$ so that the pore pressure p increases from hydrostatic at the gouge country rock

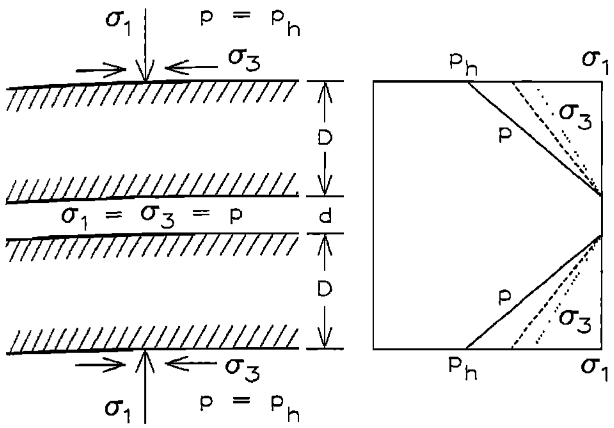

Fig. 2. A schematic diagram showing how σ_1, σ_3 and P varies within a fault zone containing material with a threshold gradient of σ_1/D. The dashed and dotted lines are the values for σ_3 when the strength of the clay gouge is such that the ration of $\bar{\sigma}_1/\bar{\sigma}_3$ is equal to 3 and 2 respectively.

interface, to lithostatic in the central zone of width d and follows the relationship,

$$p = p_h + \frac{(\sigma_1 - p_h)}{D} X \qquad (1)$$

where X is the distance from the gouge country rock interface. If $\sigma_1 = 2.5\ p_h$ and $\bar{\sigma}_3 = \frac{1}{3}\ \bar{\sigma}_1$ in the gouge zone then by substitution,

$$\sigma_3 = 1.5\ p_h + p_h \frac{X}{D} \qquad (2)$$

which is shown as the dashed line in Figure 2. If the gouge is weaker so that $\bar{\sigma}_3 = \frac{1}{2}\ \bar{\sigma}_1$ then,

$$\sigma_3 = 1.75\ p_h + .75\ p_h \frac{X}{D} \qquad (3)$$

which is shown as the dotted line in Figure 2. In both cases $\sigma_3 = \sigma_1 = p$ in the central region.

There are three important consequences of this model that help to resolve the fault mechanics paradox previously described. First, the pore pressure in the central zone can be as high as the lithostatic pressure, and can be maintained at this level indefinitely. Second, the pore pressure never exceeds the minimum principal stress regardless of the strength of the gouge so the water cannot escape by hydrofracturing the country rock. Finally, the fault zone has zero shear strength so that movement can occur on the fault with normal or subnormal compression. Therefore, the constraints of the heat flow data and the orientation of the stress directions are satisfied.

If the TG is a function of the effective pressure, then in the central zone p may be less than σ_1 and the central zone will have a finite strength. So that in this case, for slip to occur on the fault σ_1 will make an angle of less than 90° to the plane of the fault. In other words, there will be subnormal fault zone compression. If the central zone has a very low shear strength, it will tend to be extruded up and out of the fault zone. Structures such as this are common along the San Andreas fault, for example, in the Mecca Hills area in southern California *Sylvester* [1988]. Another consequence of our model is the fact that, regardless of the orientation of the fault zone to the far field stresses the fault will yield in such a way as to force the near field maximum principal stress to rotate until it is sub–normal to the fault. This effect which is a consequence of the low strength of the fault, is similar to the stress field rotation near the surface of an open crack since, in that case, the free surface requires zero shear tractions on the crack walls. Observations of stress orientations in California are consistent with this prediction [*Zoback et al.*, 1987].

Another consequence of this model is that the minimum width of the gouge layer will depend on depth for this mechanism to operate throughout the seismogenic zone. The pressure gradient will be the difference between the pressure in the central zone σ_{1z} and the hydrostatic pressure p_{hz} at depth z divided by the width of the zone D_z. If the TG is 0.5 MPa/m independent of temperature, effective pressure and composition then

$$T_z \geq 2D_z = 2(\sigma_{1z} - p_{hz})/TG \qquad (4)$$

where T_z is the thickness of the gouge zone. If the density of the overlying rocks is 2,500 Kg/m^3 then T_z = 60 m/km. At the surface, the gouge zone can have zero width and at a depth of 15 km it only needs to be 900 m wide to satisfy our model. The central zone can be any width. By measuring the reflection velocities along profiles across the San Andreas fault in the Chalome Valley region *Shedlock et al.* [1990] determined that the width of the fault zone is less than 50 m wide near the surface but it apparently widens with depth. Unfortunately, reflection data is not accurate enough to make a more precise statement than this. With subsidiary faults that are less mature than the San Andreas, it would be expected that the width of their gouge zones would be much smaller, even though the same physical mechanism may operate on them as well, they will be much stronger because the pore pressure in their fault zones cannot build up to lithostatic values. Thus, movement on secondary faults will be minor compared to what it is on the San Andreas, even though they are subjected to higher shear stress because of their more favorable orientation.

We would expect that the mechanism that we have proposed here would be more effective on the long straight section of the creeping zone in central California. In the locked regions of northern and southern California, earthquakes could still occur if most of the normal stress on the fault is supported by gouge with high pore pressure, but the fault is hindered from sliding by one or more interlocked asperities that result in an earthquake when they fail, as simulated in the laboratory experiments carried out by *Lockner and Byerlee*, [1989]. In this case σ_1 may in general be normal to the fault, but locally where the asperities interlock it may be 45° to the fault. Alternatively, in the locked regions the gouge may not be wide enough for the pore pressure in its central region to be much greater than hydrostatic and in this region the fault could be strong almost everywhere, so that σ_1 may be close to 45° to the fault everywhere. The compilation of the data by *Zoback et al.* [1987] shows that there is a large variation in the direction of the maximum horizontal stress along the San Andreas fault, but it is not clear that this is real or simply an artifact of the methods used in their determination. Thus our present confidence in the data is not sufficient to distinguish between the two possibilities set out above.

Finally, it should be pointed out that the mechanism that we have proposed here to explain the rheology of some strike–slip faults may also explain how pore pressures close to lithostatic can be maintained in fault zones, for geologically significant lengths of time to allow slip on large overthrusts and on very low angle normal faults without gross failure of the overlying rocks.

References

Fertl, W.H., Abnormal formation pressures, 382 pp., Elsevier, New York, 1976.

Hansbo, S., Consolidation of clay with special reference to the influence of vertical sand drains, 160 pp., Proceeding 18, Swedish Geotechnical Inst., Stockholm, 1960.

Hanshaw, B.B., and J.D. Bredehoeft, On the maintenance of anomalous fluid pressures: II Source layer at depth, *Geol. Soc. Am. Bull.*, 79, 1107–1122, 1968.

Homola, A.M., J.N. Israelachvili, M.L. Gee, and P.M. McGuiggan, Measurements of and relation between the adhesion and friction of two surfaces separated by molecularly thin liquid films, *Journal of Tribology*, 111, 675–682, 1989.

Israelachvili, J.N., P.M. McGuiggan and A.M. Homola, Dynamic properties of molecularly thin liquid films, *Science*, 240, 189–191, 1988.

Lockner, D., and J. Byerlee, An example of slip instability resulting from displacement varying strength, *PAGEOPH*, 133, 269–281, 1990.

Lutz, J.F., and W.D., Kemper, Intrinsic permeability of clay as affected by clay water interactions, *Soil Science*, 88, 83–90, 1959.

McGuigan, P.M., J.N. Israelachvili, M.L. Gee and A.M. Homola, Measurements of static and dynamic interactions of molecularly thin liquid films between solid surfaces, *Mat. Res. Soc. Symp.*, 140, 79–88, 1989.

Miller, R.J., and P.F. Low, Threshold gradient for water flow in clay systems, *Proc. Soil Sci. Am.*, 27, 605–609, 1963.

Ping, L.S., Measuring extremely low flow velocity of water in clays, *Soil Science*, 95, 410–413. 1963.

Radney, B., and J. Byerlee, Laboratory studies of the shear strength of montmorillonite (abstract), Eos Trans. AGU, 69, 1463, 1988.

Shedlock, K.M., T.M. Brocher, and S.T. Harding, Shallow structure and deformation along the San Andreas fault in Chalome Valley, California, based on high–resolution reflection profiling, *J. Geophys. Res.*, 95, 5003–5020, 1990.

Swartzendruber, D., Non–darcy flow behavior of liquid saturated porous media, *J. Geophys. Res.*, 67, 5205–5213, 1962.

Sylvester, A.G., Strike–slip faults, *Geol. Soc. Am. Bull.*, 100, 1666–1703, 1988.

Von Engelhardt, W., and W.L.M. Tunn, The flow of fluids through sandstones. (Translated by P.A. Witherspoon from *Heidelbeger Beitrage Zur Mineralogie und Petrographie* 2,12–25, 1954), *Illinois State Geol. Survey Cir.*, 194, 1–16, 1955.

Watts, N.L., Theoretical aspects of cap–rock fluid seals for single and two phase hydrocarbon columns, *Marine and Petroleum Geology*, 4, 274–307, 1987.

Zoback, M.D., and J. Healy, Overview of In–situ stress measurements in the Cajon Pass scientific drill hole to a depth of 3.5 km, (abstract) Eos Trans. AGU, 70, 480, 1989.

Zoback, M.D., M.L., Zoback, V.S. Mount, J. Suppe, J.P. Eaton, J.H. Healy, D. Oppenheimer, P. Reasenberg, L. Jones, C.B. Raleigh, I.G. Wong, O. Scotti, and C. Wentworth, New evidence on the state of stress of the San Andreas fault system, *Science*, 238, 1105–1111, 1987.

J. Byerlee (U.S. Geological Survey, Menlo Park, CA 94025)

(Received: July 13, 1990
Revised: September 6, 1990
accepted: September 11, 1990)