U.S. DEPARTMENT OF THE INTERIOR

U.S. GEOLOGICAL SURVEY

# PROCEEDINGS OF

# WORKSHOP LXIII

# THE MECHANICAL INVOLVEMENT OF FLUIDS IN FAULTING

6 - 10 JUNE 1993

Sponsored by

U.S. GEOLOGICAL SURVEY

NATIONAL EARTHQUAKE HAZARDS REDUCTION PROGRAM

Editors and Convenors

Stephen Hickman
U.S. Geological Survey
Menlo Park, California 94025-3591

Richard Sibson
University of Otago
Dunedin, New Zealand

Ronald Bruhn
University of Utah
Salt Lake City, Utah 84112

OPEN-FILE REPORT 94-228

Compiled by
Muriel Jacobson



MENLO PARK, CALIFORNIA
1994

## EVIDENCE FOR FLUID INVOLVEMENT IN FAULTING AND DEEP CRUSTAL FLUID RESERVOIRS

FAULT-ZONE TRANSPORT PROPERTIES AND COMPOSITION OF FAULT-ZONE FLUIDS

COUPLED MECHANICAL AND HYDROLOGICAL PROCESSES IN FAULTING

CHEMICAL EFFECTS OF FLUIDS ON FAULT-ZONE RHEOLOGY

## ACKNOWLEDGMENTS

The success of the Red Book Conference on the Mechanical Involvement of Fluids in Faulting was due, in no small part, to the skill and hard work of Susan Larsen and Barbara Simpson in making most of the logistical and travel arrangements for this meeting. We are also indebted to Muriel Jacobson, Barbara Simpson and Rose Trombley for their considerable help and expertise in producing and distributing this Red Book volume and to Wanda Seiders for locating potential conference facilities in the Sierra Nevada foothills. Discussions with Stephen Kirby, John Logan, Walter Mooney, Elaine Padovani and David Schwartz were influential in deciding on both the format and the invitation list for this conference Thanks also to Stephen Cox, Richard Henley, John Logan, Evelyn Roeloffs and John Rudnicki for helping moderate the group discussions and to James Elliot for assisting with the field trip.

This Red Book Conference was sponsored by the National Earthquake Hazards Reduction Program of the United States Geological Survey (USGS). Additional support for publication of the Red Book volume was provided by the USGS Deep Continental Studies Program.

Stephen Hickman
Richard Sibson
Ronald Bruhn

# LIST OF PARTICIPANTS

Dr. Yehuda Ben-Zion
Harvard University
Dept. of Earth & Planetary Sciences
24 Oxford Street
Cambridge, MA 02138

Dr. Michael Blanpied
U.S. Geological Survey
345 Middlefield Rd., MS977
Menlo Park, CA 94025

Dr. Sue Brantley
College of Earth & Mineral Sciences
Earth System Science Center
Penn State University
University Park, PA 16802-2711

Dr. Stephen Brown
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185

Dr. Ronald Bruhn
Dept. of Geology & Geophysics
University of Utah
717 W.C. Browning Bldg.
Salt Lake City, UT 84112-1183

Dr. James Brune
University of Nevada, Reno
Dept. of Geological Sciences
Reno, NV 89557-0141

Dr. James Byerlee
U.S. Geological Survey
345 Middlefield Rd., MS 977
Menlo Park, CA 94025

Dr. Fred Chester
Saint Louis University
Dept. of Earth & Atmospheric Sciences
3507 Laclede Ave.
St. Louis, MO 63103

Dr. Stephen Cox
Australian National University
Research School of Earth Sciences
GPO Box 4
Canberra ACT 2601
Australia

Dr. Daniel Davis
Dept. of Earth & Space Sciences
State University of New York
Stony Brook, NY 11794-2100

Dr. Thomas Dewers
School of Geology & Geophysics
University of Oklahoma
810 Sarkey's Energy Center, Suite 810
Norman, OK 73019

Dr. Donna Eberhart-Phillips
Institute of Geologic & Nuclear Sciences
Kelburn Office
P.O. Box 1320
Wellington, New Zealand

Mr. James Elliott
Dept. of Geology
University California
Santa Barbara, CA 93106

Dr. Terry Engelder
Dept. of Geosciences
336 Deike Bldg.
Penn State University
University Park, PA 16802

Dr. Brian Evans
Dept. of Earth, Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Dr. James Evans
Utah State University
Dept. of Geology
Logan, UT 84322-4505

Dr. Craig Forster
Dept. of Geology & Geophysics
719 W.C. Browning Bldg.
University of Utah
Salt Lake City, UT 84112-1183

Dr. Jean-Pierre Gratier
LGIT-CNRS (URA 733)
Universite Joseph Fourier
IRIGM
BP53X
38041 Grenoble, France

Dr. Richard W. Henley
Etheridge & Henley, Geoscience Consultants
P.O. Box 3778
Manuka ACT 2603
Australia

Dr. Stephen Hickman
U.S. Geological Survey
345 Middlefield Rd., MS 977
Menlo Park, CA 94025

Dr. David Hill
U.S. Geological Survey
345 Middlefield Rd., MS 977
Menlo Park, CA 94025

Dr. Roy Hyndman
Pacific Geoscience Center
P.O. Box 6000
Sidney BC V8L4B2
Canada

Dr. Peggy Johnson
University of California, Berkeley
Dept. of Geology and Geophysics
Berkeley, CA 94720

Dr. Malcolm Johnston
U.S. Geological Survey
345 Middlefield Rd., MS 977
Menlo Park, CA 94025

Dr. Robert Kerrich
University of Saskatchewan
Dept. of Geological Sciences
Saskatoon SK S7N 0W0
Canada

Dr. Geoffrey King
Institut de Physique du Globe
5 Rue Rene Descartes
67084 Strasbourg Cedex
France

Dr. Rob Knipe
Dept. of Earth Sciences
The University of Leeds
Leeds LS2 9JT
England

Ms. Susan Larsen
U.S. Geological Survey
345 Middlefield Rd., MS 977
Menlo Park, CA 94025

Dr. John Logan
Center for Tectonophysics
Texas A & M University
College Station, TX 77843-3113

Dr. Peter Malin
Dept. of Geology
Box 90235
Duke University
Durham, NC 27708-0235

Dr. J. Casey Moore
University of California, Santa Cruz
Earth Sciences Board of Studies
Santa Cruz, CA 95064

Dr. Craig Nicholson
Institute for Crustal Studies
University of California
Santa Barbara, CA 93106

Dr. Elaine Padovani
U.S. Geological Survey
Office of Earthquakes, Volcanoes & Engineering
905 National Center
Reston, VA 22092

Dr. Evelyn Roeloffs
U.S. Geological Survey
Cascades Volcano Observatory
5400 MacArthur Blvd.
Vancouver, WA 98661

Dr. Stuart Rojstaczer
Duke University
Dept. of Geology
206 Old Chemistry Bldg.
Durham NC 27706

Dr. Allan Rubin
Princeton University
Dept. of Geology &
Geological Sciences
Princeton, NJ 08544

Dr. John Rudnicki
Northwestern University
Dept. of Civil Engineering
2145 Sheridan Road
Evanston, IL 60208-3109

Dr. Christopher H. Scholz
Lamont-Doherty Earth Observatory
Palisades, NY 10964

Dr. Paul Segall
Stanford University
Dept. of Geophysics
Stanford, CA 94305-2215

Dr. Rick Sibson
Dept. of Geology
University of Otago
P.O. Box 56
Dunedin
New Zealand

Dr. Terry E. Tullis
Dept. of Geological Sciences
Box 1846
Brown University
Providence, RI 02912

Dr. Phillip Wanamaker
University of Utah
Research Institute
391 Chipeta Way, Suite C
Salt Lake City UT 84108-1295

Dr. Bob Wintsch
Indiana University, Bloomington
Dept. of Geological Sciences
1005 E. 10th Street
Bloomington IN 47405-5101

Dr. Mark Zoback
Stanford University
Dept. of Geophysics
Stanford CA 94305-2215

# INTRODUCTION: USGS RED-BOOK CONFERENCE ON THE MECHANICAL INVOLVEMENT OF FLUIDS IN FAULTING

**Stephen Hickman[1], Richard Sibson[2], and Ronald Bruhn[3]**

[1]U.S. Geological Survey, 345 Middlefield Rd., MS 977, Menlo Park, CA 94025
[2]Department of Geology, University of Otago, P.O. Box 56, Dunedin, New Zealand
[3]University of Utah, Department of Geology and Geophysics, Salt Lake City, UT 84112

A tantalizing body of evidence suggests that fluids are intimately linked to a variety of faulting processes, including the nucleation, propagation, arrest and recurrence of earthquake ruptures; fault creep; and the long-term structural and compositional evolution of fault zones. Beyond the widely recognized physical role of fluid pressures in controlling the strength of crustal fault zones, it is also apparent that fluids can exert mechanical influence through a variety of chemical effects.

A "Red-Book" Conference on the Mechanical Effects of Fluids in Faulting was sponsored by the United States Geological Survey under the auspices of the National Earthquake Hazards Reduction Program (NEHRP) at Fish Camp, California, from June 6–10, 1993. The purpose of the conference was to: 1) draw together and evaluate the disparate evidence for the involvement of fluids in faulting; 2) establish communication on the importance of fluids in the mechanics of faulting between the different disciplines concerned with fault-zone processes; and 3) help define future critical investigations, experiments and observational procedures for evaluating the role of fluids in faulting. Some of the questions addressed at this workshop included:

1) What are fluid pressures at different levels within seismically active fault zones? Do they remain hydrostatic throughout the full depth extent of the seismogenic regime, or are they generally superhydrostatic at depths only in excess of a few kilometers?

2) Are fluid pressures at depth within fault zones constant, or are they time-dependent? What is the expected spatial variability in fluid pressures?

3) What is the role of crustal fluids in the overall process of stress accumulation, release and transfer during the earthquake cycle? Through what mechanisms might fluid pressure act to control the processes of rupture nucleation, propagation and arrest?

4) What is the chemical role of fluids in facilitating fault creep, including their role in aiding solid-state creep and brittle fracture processes and in facilitating solution-transport deformation mechanisms?

5) What are the chemical effects of aqueous fluids on constitutive response, frictional stability and long-term fault strength?

6) What are the compositions and physical properties of fault fluids at different crustal levels?

7) What are the mechanisms by which porosity and permeability are either created or destroyed in the mid to lower crust? What factors control the rates of these processes? How should these effects be incorporated into models of time-dependent fluid transport in fault zones?

8) What roles do faults play in distributing fluids in the crust and in altering pressure domains? In other words, when and by what mechanisms do faults aid in or inhibit fluid migration? What are the typical fluid/rock ratios, flow rates and discharges for fault zones acting as fluid conduits?

9) Are fluids present in the subseismogenic crust, and by what transformation and/or transport processes are they incorporated into the shallower seismogenic portions of faults?

A diverse group of 45 scientists attended this conference, including researchers on electrical and magnetic methods, geochemistry, hydrology, ore deposits, rock mechanics, seismology and structural geology. To provoke as much interdisciplinary exchange as possible, the conference was divided into four plenary sessions, each session consisting of about 10–13 poster presentations followed by a moderated group discussion. These sessions were: 1) Evidence for Fluid Involvement in Faulting and Deep Crustal Fluid Reservoirs, 2) Fault-Zone Transport Properties and Composition of Fault-Zone Fluids, 3) Coupled Mechanical and Hydrological Processes in Faulting and 4) Chemical Effects of Fluids on Fault-Zone Rheology. At the beginning of each session, we asked each person presenting a poster in that session to give a brief (2–3 minute) synopsis of their poster presentation. At the conclusion of each session, the group reconvened for about one hour to discuss issues raised during that session. Very lively discussions centered on field evidence for or against super-hydrostatic fluid pressures along active faults, the origin and nature of fluid pressure compartments and their role in triggering earthquakes, and the nature of hydrologic perturbations induced by earthquakes. On the second day of the conference, Rick Sibson led a one-day field trip through the mesothermal gold-bearing quartz-vein systems of the Sierran foothills (the *Mother Lode* country), where participants saw and discussed evidence for episodic fluid flow driven by high fluid pressures in the exhumed reverse-slip Melones fault zone. The conference concluded with small-group discussions in which participants talked about the major paradigms and unresolved questions pertaining to fluids in faulting and identified future research needed to help answer these questions. Summaries of these small-group discussions were presented to the conference participants in a final "wrap-up" session.

Workshop participants identified a number of key areas for future research pertaining to the mechanical involvement of fluids in faulting, each of which requires integration of knowledge from several fields and communication between specialists with different backgrounds. The primary goal of these studies is to identify those processes and parameters which are most important in controlling fault-zone rheology and which will, therefore, dominate the mechanics of faulting at different levels in the crust. These research topics fall into three broad categories: 1) geological observations of exhumed fault zones, 2) laboratory experiments on natural and synthetic fault-zone materials and 3) fault-zone drilling combined with surface-based geophysical and geological investigations.

Studies of exhumed fault zones must be carefully designed to provide information on deformation mechanisms that operate at different crustal levels, together with fluid-inclusion and mineral-equilibria analyses that constrain the compositions, pressures and temperatures of fluids both within and adjacent to fault zones. As noted by several conference participants, stress heterogeneities induced by fault slip and deformation can lead to considerable uncertainties in inferring past fluid pressures simply from observations of vein geometry in outcrop. Thus, structural and fluid-inclusion/mineralogical techniques should be combined to constrain fluid pressures during vein formation at different stages in the faulting cycle. Particular attention should also be paid to evidence for fluid communication, or lack thereof, between fault zones and country rock and evidence for or against cyclic dilatancy and gouge compaction within fault zones. Careful mapping is needed to document the spatial scales over which structural, mineralogical, fluid-chemical and hydraulic heterogeneity exists. More data on the existence, origin, dimensions and temporal characteristics of abnormally pressured fluid compartments within and adjacent to fault zones are of paramount importance in this regard. In all of these investigations, careful microstructural studies and determinations of past temperature and pressure conditions are required to assess the degree of annealing and other secondary alteration on samples collected from exhumed fault zones, where a complex history of uplift and denudation may have severely altered, or even destroyed, evidence for deformation mechanisms operative during fault slip.

Although it is generally accepted that mylonites with well-ordered fabrics are predominantly the result of aseismic plastic shearing, with the exception of localized melts generated by rapid seismic

slip (i.e., the pseudotachylytes occasionally found in exhumed fault zones) there is currently no reliable way to distinguish the cataclastic products of seismic versus aseismic slip in fault zones. New laboratory friction experiments are needed at high rupture-propagation and sliding velocities to constrain the mechanisms of dynamic rupture and, perhaps, identify microstructures diagnostic of slip speed and stability. More laboratory experimentation is also required to document the importance of solution-transport reactions and other fluid-rock interactions in controlling long-term fault strength, the stability of sliding, the evolution of fluid pressures along faults, and the time scales for interseismic strength recovery in the mid to lower crust. Carefully controlled experiments to determine the rates at which various deformation and sealing/strengthening processes operate—and how these rates depend upon mineralogy, fluid chemistry, grain size and pore geometry—is considered to be one of the most crucial goals for experimentalists. Finally, the importance of electrokinesis as a potential earthquake precursor was noted, together with the need for experiments that investigate electromagnetic properties of rocks during fluid flow at elevated temperature and pressure.

Large-scale field experiments using geophysical imaging techniques combined with borehole observations are urgently needed to confirm interpretations and hypotheses related to the mechanical role of fluids in faulting arising from observations on exhumed faults and laboratory experiments. Drilling provides the only direct means of measuring in-situ stress, fluid pressure and permeability; collecting fluid and rock samples from active faults and wall rocks; and monitoring time-dependent changes in fluid pressure and chemistry, deformation and stress during the earthquake cycle. Drilling and downhole measurements in and adjacent to active fault zones must be accompanied by detailed surface geophysical and geological observations to allow extrapolation of fluid pressure regimes and fault-zone physical properties over a much larger volume than sampled by the borehole. For example, seismic tomography coupled with borehole monitoring may provide one of the best methods for real-time monitoring of temporal changes in fluid pressure and rock physical properties within and adjacent to fault zones during earthquakes and episodic fault slip.

Virtually all of the workshop participants are represented by papers in this Red Book volume, based, in large part, on their individual poster presentations. These papers are organized into four separate chapters, corresponding to the four workshop sessions identified above. It should be noted, however, that in cases where papers dealt with topics relevant to two or more sessions these chapter assignments were somewhat arbitrary. About 20 of these papers will be appearing in a special issue of the *Journal of Geophysical Research* to be published early in 1995.

# 1: EVIDENCE FOR FLUID INVOLVEMENT IN FAULTING AND DEEP CRUSTAL FLUID RESERVOIRS

# DEFORMATION PROCESSES AND CONTROLS ON FAULT MECHANICS DURING FAULT-VALVE BEHAVIOR IN THE WATTLE GULLY FAULT ZONE, CENTRAL VICTORIA, AUSTRALIA

Stephen F. Cox

Research School of Earth Sciences, The Australian National University, Canberra, ACT 0200, Australia

## ABSTRACT

The Wattle Gully Fault is a high-angle reverse fault zone that developed at depths around 10km late during mid-Devonian regional crustal shortening in central Victoria. The fault has a length around 800m and a maximum displacement of 40m. The widespread development of steeply-dipping fault-fill quartz veins and associated sub-horizontal extension veins within the fault zone indicate that faulting occurred at low stress differences in a transiently supralithostatic fluid pressure regime. The structure was the site of focussed fluid migration during vein formation.

The internal structures of veins indicate vein development was intimately associated with faulting and involved thousands of increments of hydraulic fracturing, fault slip and vein sealing. Fault slip and vein opening is interpreted to have been strongly influenced by repeated fluid pressure fluctuations associated with fault-valve behavior.

The geometries and internal structures of veins in the Wattle Gully Fault Zone are interpreted in terms of cyclic changes in shear stresses, fluid pressures, and near-field principal stress orientations during fault-valve behavior. Substantial fluctuations in fluid pressure have played a key role in controlling fault strength, as well as nucleation and recurrence of slip events. There is evidence that faulting in supralithostatic fluid pressure regimes can result in near-total stress relief and local dynamic overshoot.

Fault-valve behavior is shown to have important implications for coupling between fault dynamics and the dynamics of fluid circulation around active faults that are sites of focussed fluid migration. In particular, fault-valve action is expected to lead to distinctly different fluid migration patterns adjacent to the fault zone before, and immediately after fault rupture.

## INTRODUCTION

Earthquakes nucleate primarily at depths between 5 and 20 km, and at temperatures between 100°C and 350°C in the Earth's continental crust (Sibson, 1983; Scholz, 1990). There is increasing evidence that within this seismogenic regime, active faults can play a crucial role in localising crustal fluid migration (Sibson et al, 1988; Cox et al, 1991a), and that fluids migrating through faults probably have a substantial influence on fault mechanics (Sibson, 1989, 1992; Evans, 1990; Chester et al, 1993). Studies of the internal structure of ancient fault zones that have been sites of intense fluid activity during seismogenesis have considerable potential to yield insights about earthquake mechanics.

This paper examines the geometry and internal structure of a relatively small, high-angle reverse fault zone that was active late during a period of major crustal shortening and regional metamorphism. The presence of laterally extensive vein systems, hydrothermal alteration, together with associated stable isotope signatures and fluid inclusion evidence, all

indicate that faulting was associated with substantial fluid-rock interaction (Cox et al, 1991a, in press). It is demonstrated that deformation in the Wattle Gully Fault Zone has involved fault-valve behavior in a transiently supralithostatic fluid pressure regime. The internal structures of the fault zone are used to provide insights into deformation processes, fluid pressure histories, and controls on fluid migration patterns that have operated during fault-valve behavior at seismogenic depths. The study also highlights the importance of coupling between fluctuations in fluid pressure and shear stress in controlling fault slip and slip recurrence in situations where faults are sites of focussed fluid flow.

## GENERAL GEOLOGY

The Wattle Gully Fault Zone (WGFZ) forms part of a network of reverse faults that developed late during regional mid-Devonian folding of an Ordovician sandstone-slate sequence in the Lachlan Fold Belt of central Victoria, Australia (Cox et al., 1991b). The fault zone hosts extensive quartz veining and associated gold mineralisation, and is inferred to have been a site of focussed fluid migration (Cox et al., 1991a, in press). The high fluid flux is interpreted to have been associated with lower crustal metamorphism and devolatilisation late during crustal thickening that was associated with regional deformation (Cox et al., 1991b).

Hydrothermal alteration styles and vein mineralogy, together with stable isotope and fluid inclusion data, indicate that faulting occurred at temperatures around 300°C. Reasonable geothermal gradients during regional metamorphism require that faulting is unlikely to have occurred at depths much shallower than around 6 to 10 km.

## FAULT GEOMETRY

The WGFZ has a maximum reverse displacement of 40 meters; it extends along strike for approximately 800 meters and down dip for at least 600 meters. Fault geometry has been influenced substantially by the pre-existing fold geometry (Fig. 1). At depth, the fault is bedding-concordant on the steeply west-dipping limb of an upright anticline. Displacement and associated deformation in this region has been restricted to a zone which is usually less than 50 cm wide. At higher structural levels, the fault intersects the anticlinal hinge zone and becomes a less steeply-dipping, bedding-discordant structure that traverses the east-dipping fold limb as a 30 meter wide zone of deformation in which displacement has occurred on numerous fault surfaces. At its intersection with the next west-dipping fold limb to the east, the fault zone again becomes a narrow, steeply-dipping, bedding-concordant structure. The non-planar geometry of the fault zone and control of the overall slip geometry by the orientation of the steeply-dipping, bedding-concordant fault segments, has led to the less steeply-dipping, bedding-discordant fault segment forming a dilatant jog which hosts over 1.3 Mt of vein quartz.

## VEIN GEOMETRY

Quartz veins occur as fault-fill structures (Fig. 2) predominantly in the bedding-discordant fault segment, and to a lesser extent in the bedding-concordant fault segments. Veins also occur as large extension vein arrays adjacent to faults in the bedding-discordant segment of the WGFZ (Fig. 3). Individual extension veins are usually subhorizontal, lenticular structures which extend along strike for distances up to about 10 meters and typically have thicknesses up to 30 cm. More irregularly-shaped, branching veins are also present. Extension veins are most abundant adjacent to individual faults in the dilatant jog segment of the WGFZ and systematically decrease in abundance and width away from faults. Extension veins are prominent over a maximum cross-strike width of around 50 meters in the jog segment.

The kinematics of the WGFZ and the overall vein geometry indicate that faulting and vein formation has occurred in a stress regime in which the far-field maximum principal stress ($\sigma_1$) was east-west directed and the far-field minimum principal stress ($\sigma_3$) was approximately vertical. The overall fault orientation was therefore inclined at $60^o$ - $70^o$ to the far-field $\sigma_1$.

The occurrence of steeply-dipping and irregularly oriented extension vein arrays, together with the more usual sub-horizontal extension veins in the dilatant jog segment, reflects the influence of local, and presumably transient changes in the stress regime during faulting. Mechanical anisotropy of the host rocks and vein rotation, due to localised high shear strain near some faults, has also influenced the geometry of extension veins. However, localised curvature of some veins immediately adjacent to faults in the dilatant jog cannot be ascribed to increasing shear strain adjacent to faults (Fig. 2b), and may indicate rotation of stress trajectories by as much as $90^o$ adjacent to some faults.

## INTERNAL STRUCTURES OF VEINS

*Extension Veins*

Extension veins are composed predominantly of quartz. They usually have fibrous to massive internal structures, but crustiform and idiomorphic fabrics also occur. Formation of fibrous to massive extension veins has involved crack-seal growth mechanisms, with individual growth increments ranging from 20μm to 100μm in thickness. Vein growth has therefore involved hundreds to thousands of crack-seal increments.

Syntaxial and stretched-crystal types of crack-seal microstructures are present in veins within sandstone units, whereas antitaxial microstructures are more usual in slate-hosted veins. Crustiform and idiomorphic vein fabrics occur especially in some carbonate-rich veins. The latter fabrics indicate a more simple, accretionary growth history that has involved only a single stage of fracture opening and sealing. The single stage growth veins have similar orientations to the crack-seal veins, indicating that the formation of the two vein types was most likely broadly contemporaneous.

The delicate structures of rhythmically repeated crack-seal inclusion bands and high crack aspect ratios ($>10^4$) for individual crack increments indicate that the lenticular extension veins have opened by repeated sub-critical crack growth in a regime of low driving pressures (Cox, 1991). The single stage growth veins probably have formed in lower aspect ratio cracks, and indicate higher driving pressures than the associated crack-seal veins. The association of the two vein types is interpreted to reflect temporal and spatial variations in driving pressures during the deformation history of the fault zone.

*Fault-Fill Veins*

The microstructures in fault-fill veins indicate that faults have been subject to a protracted deformation history that has involved repeated cycles of fault dilation, hydrothermal sealing and reverse slip, as well as shortening at high angles to fault surfaces. Fault slip has probably involved both aseismic and seismic slip mechanisms.

Primary crack-seal and crustification microstructures in fault-fill veins are overprinted to varying degrees by the effects of both brittle deformation and dissolution-precipitation processes. Plastic deformation has been locally important. Discrete zones of pervasive shear failure are marked by abrupt truncation of internal vein structures and the occurrence of narrow bands of fine-grained quartz cataclasite in otherwise coarse-grained vein quartz. Localised slip along the margins of fault-fill veins and internal wall-rock screens is indicated by the presence of slaty cataclasites and breccias. Intense stylolitisation has also occurred at the margins of many fault-fill veins and along the edges of internal wall-rock screens. The occurrence of undeformed and strongly deformed vein fabrics in adjacent laminae of laminated fault-fill veins, together with overprinting relations between these various types of microstructures, indicate that dilation, slip, and shortening at high angles to the fault, have been part of a cyclic process.

## DISCUSSION

*Fluid Pressures and Stress Regimes*

The widespread development of sub-horizontal extension veins during high-angle reverse faulting in the WGFZ indicates that faulting has occurred in a low stress difference regime in which fluid pressures were at least transiently supralithostatic. Evidence for repeated opening and sealing of extension veins requires that fluid pressures have fluctuated between supralithostatic and sub-lithostatic levels during faulting. Clear evidence of repeated episodes of dilatation and slip on the steeply-dipping, and therefore severely misoriented fault segments, also requires supralithostatic fluid pressures and low fault strength (Sibson, 1990).

The internal structures of the WGFZ provide valuable insights about the way in which the mechanics of fault zones can depend on interactions between hydrothermal sealing of faults and fluctuations in fluid pressure and shear stress during the seismic cycle. In particular, slip recurrence intervals and the time-dependence of fault shear strength will be particularly influenced by the relative rates of reaccumulation of fluid pressure and shear stress after episodes of fault failure (Fig. 4).

The presence of extension veins adjacent to faults indicates that hydrothermal self-sealing and formation of fault-valves within the WGFZ after successive shear failure events commonly has been sufficiently rapid that supralithostatic fluid pressures have developed before shear stresses could recover to levels high enough to induce renewed shear failure (Fig. 5). The occurrence of irregularly-oriented vein stockworks in several parts of the dilatant jog fault segment, and evidence for repeated dilatation of the steeply-dipping fault segments in the WGFZ is especially significant as it indicates that supralithostatic fluid pressures have been attained locally in an isostatic or near-isostatic stress regime. This requires that some failure events have led locally to total shear stress relief.

The localised occurrence of steeply-dipping extension veins is inconsistent with the far-field stress regime inferred for reverse faulting in the WGFZ. A transient change in the orientation of the near-field $\sigma_1$ from its usual sub-horizontal attitude to a sub-vertical orientation, as required by this vein geometry, may have a number of causes. One possibility is that some failure events have resulted in dynamic overshoot and a transient flip in the orientation of the stress-field. Alternatively, local stress reorientations can occur due to interactions between adjacent fault segments. Transient near-field stress re-orientation can also develop in response to wall-rock transport over non-planar fault geometries during slip episodes. Such localised stress re-orientation is interpreted to indicate that far-field stress differences have been at least transiently low during some stages in the evolution of the WGFZ.

Comparison of the maximum rupture dimensions of the WGFZ with modern earthquake ruptures (Sibson, 1989) indicates that individual slip increments have likely been up to several centimeters. This corresponds to earthquakes in the M3 to M4 range and indicates that the total displacement on the WGFZ could have involved several thousand slip events. This number of slip events is comparable to the number of crack-seal growth increments involved in the formation of the larger extension veins adjacent to the fault zone, and lends support to the concept that crack-seal growth cycles correlate with individual fault-valve cycles.

*Fluid Dynamics During Fault-Valve Behavior in the Wattle Gully Fault Zone*

Fault-valve behavior may have important implications for the dynamics of fluid circulation and hydrothermal alteration in active fault zones, such as the WGFZ, which are sites of focussed crustal fluid outflow.

Prior to failure events, upward fluid migration is impeded when low permeability regions form within the fault zone by localised hydrothermal sealing. This promotes fluid infiltration of the wall-rocks below pressure seals, especially when increased fluid pressures

enhance hydrofracture dilatancy adjacent to faults (Fig. 6). Breaching of seals during fault failure can interrupt this flow regime by transiently increasing fault permeability and thereby causing an abrupt decrease in fluid pressure within the fault zone. Fluids that have reacted with wall-rocks can then be driven back into the transiently lower fluid pressure fault zone. Such effects are expected to be particularly important in the vicinity of dilatant jogs.

Fluid inclusion data so far provide only limited constraints on the magnitude of fluid pressure fluctuations associated with fault-valve action in the WGFZ. Low salinity aqueous fluid inclusions, from vuggy fault-fill quartz in the dilatant jog segment, homogenise in the range 160°C to 240°C (Cox et al, in press). Provided trapping occurred at around 300°C, as indicated by alteration mineralogy and stable isotope data, the 80°C spread in homogenisation temperatures allows a maximum range in trapping pressures between 80 MPa and 175 MPa. However, some of the variation in homogenisation temperatures is likely to reflect real variations in trapping temperatures as well as trapping pressures. Substantial variations in the density of $CO_2$ (0.4 - 0.85 gm $cm^{-3}$) in liquid $CO_2$ -bearing fluid inclusions also support the concept that fluid pressures have varied during faulting and associated vein formation. However, it is not clear whether these changes have been cyclic or monotonic during the evolution of the fault zone.

In view of the arguments for fluctuations in fluid pressures in the WGFZ, it is expected that fault-valve action could lead to distinctly different fluid migration patterns before, and immediately after fault rupture (Fig. 7). Cyclic episodes of pre-failure fluid-rock reaction adjacent to the fault zone, followed by immediate post-failure fluid mixing, during fault-valve action, appear to have been key factors controlling the evolution of fluid chemistry and gold deposition in the WGFZ (Cox et al, in press).

*Fluid Volumes*

Analysis of fluid budgets on the basis of the mass of quartz and gold deposited in the WGFZ requires that a minimum of several cubic kilometers of water has been channeled up through the structure during fault-valve activity (Cox et al, 1991a). This requires approximately $10^6 m^3$ of fluid to have migrated up through the fault during each of several thousand fault-valve cycles.

For reasonable lifetimes of the hydrothermal system ($10^4$ - $10^5$ years), and assuming that near-lithostatic vertical fluid pressure gradients have been maintained, such flow rates require a time-averaged permeability for the fault zone in the range $10^{-12}$ - $10^{-14}$ $m^2$. Such high values presumably reflect the transiently rapid fluid discharge through the fault zone in the immediate post-failure part of the fault-valve cycle. The development of supralithostatic fluid pressure compartments, albeit "leaky" ones, in response to hydrothermal self-sealing and formation of fault-valves in the interseismic periods, requires a very much lower fault permeability, probably less than $10^{-20}$ $m^2$.

## CONCLUSIONS

The Wattle Gully Fault Zone has been a site of localised crustal fluid outflow late during a period of regional crustal shortening and low grade metamorphism. The internal structures of this reverse fault indicate that its evolution has involved repeated episodes of fault slip, fault dilatation, and growth of spatially associated extension fractures. Hydrothermal self-sealing of the fault zone after repeated failure events has promoted fault-valve behavior and development of transiently supralithostatic fluid pressures. Rates of recovery of fluid pressure and shear stress after failure events are interpreted to been important factors influencing nucleation and recurrence of failure events.

The abundance of subhorizontal hydraulic extension fractures, evidence for repeated dilatation of steeply-dipping fault segments, and the severe misorientation of the WGFZ with respect to the far-field maximum principal stress, all indicate that repeated fault failure has occurred at supralithostatic fluid pressures and relatively low stress differences. The

geometries of extension veins in the fault zone indicate that some failure events have been associated with near-total relief of shear stress and transient re-orientation of near-field stresses.

Repeated fluctuations in fluid pressure associated with fault-valve behavior are expected to lead to distinctly different fluid migration patterns adjacent to faults before, and immediately after rupture events.

## Acknowledgements

M. Etheridge, R. Henley, S. Matthai, T. Potter, R. Sibson, and J. Streit are thanked for many thought-provoking discussions over a number of years. This work has been supported by an ARC Senior Reseach Fellowship. USGS generously provided financial support to allow the author to participate in the Fish Camp workshop.

## References

Chester, F.M, Evans, J.P. and Biegel, R.L., 1993. Internal structure and weakening mechanisms of the San Andreas Fault. *J. Geophys. Res.* 98, 771 - 786.

Cox, S.F., 1991. Geometry and internal structures of mesothermal vein systems: Implications for hydrodynamics and ore genesis during deformation. In: "Structural Geology in Mining and Exploration". Extended Abstracts. Geology Department (Key Centre) & University Extension, The University of Western Australia, Publication No. **25**, 47-53.

Cox, S.F., Wall, V.J., Etheridge, M.A. & Potter, T.F., 1991a. Deformational and metamorphic processes in the formation of mesothermal vein-hosted gold deposits - Examples from the Lachlan Fold Belt in central Victoria, Australia. *Ore Geology Reviews* **6**, 391-423.

Cox, S.F., Etheridge, M.A., Cas, R.A.F. & Clifford, B.A., 1991b. Deformational style of the Castlemaine area, Bendigo-Ballarat Zone: Implications for evolution of crustal structure in central Victoria. *Australian Journal of Earth Sciences* **38**, 151-170.

Cox, S.F. Sun, S.-S., Etheridge, M.A., Wall,V.J. and Potter, T.F., in press. Structural and geochemical controls on the development of turbidite-hosted gold-quartz vein deposits, Wattle Gully Mine, central Victoria, Australia. *Economic Geology.*

Evans, J.P., 1990. Textures, deformation mechanisms, and the role of fluids in the cataclastic deformation of granitic rocks. In: Knipe, R.J. and Rutter, E.H. (eds), "Deformation Mechanisms, Rheology and Tectonics". *Geol. Soc Special Publn.* 54, pp 29 -39.

Scholz, C.H., 1990. The Mechanics of Earthquakes and Faulting. Cambridge University Press, Cambridge, 439 pp.

Sibson, R.H, 1983. Continental fault structure and the shallow earthquake source. *J. Geol. Soc. London* **140**, 741 - 767.

Sibson, R.H., 1989. Earthquake faulting as a structural process. *Journal of Structural Geology* **11**, 1-14.

Sibson, R. H., 1990. Rupture nucleation on unfavorably oriented faults. *Bull. Seism. Soc. Am.* **80**, 1580-1604.

Sibson, R. H., 1992. Implications of fault-valve behaviour for rupture nucleation. *Tectonophysics* **211**, 283 - 293.

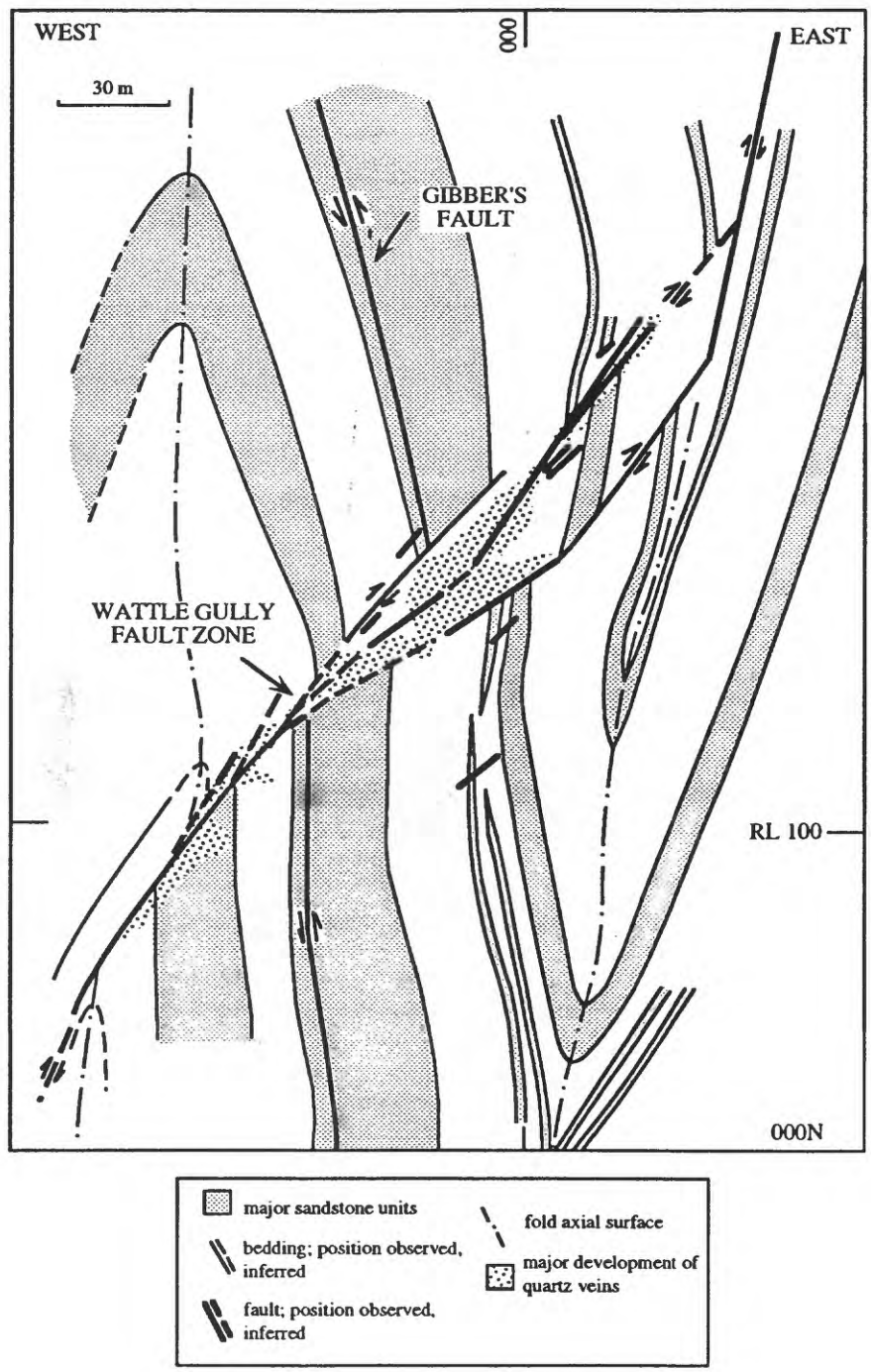



**Figure 1.** Structure of the Wattle Gully Fault Zone in an east-west section through the main shaft.

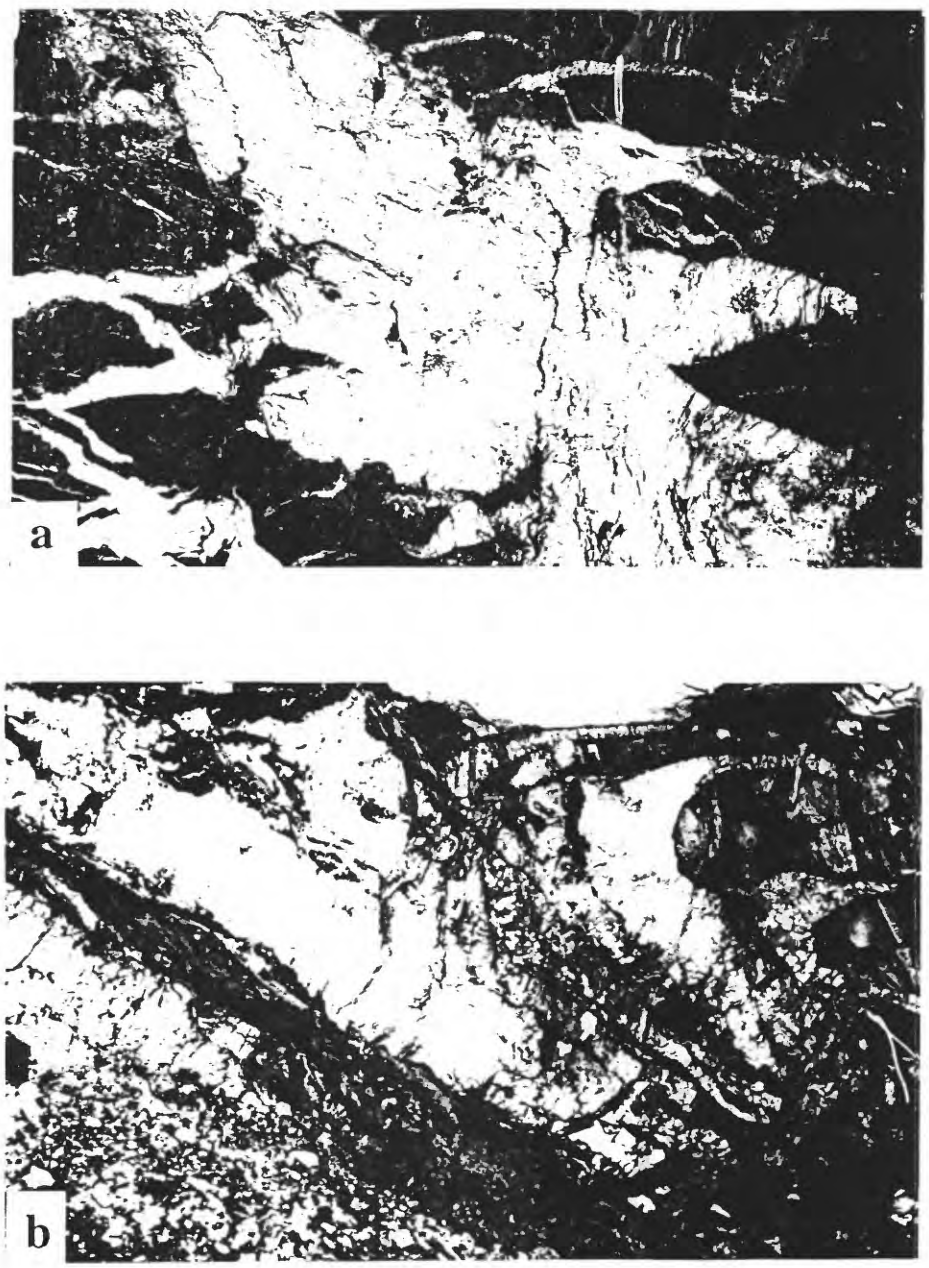



**Figure 2.**
**a.** Fault-fill quartz vein and associated extension veins in a minor reverse fault zone. The field of view is about 2 meters wide.
**b.** Fault-fill veins and associated extension veins in the central part of the WGFZ. Note that some of the fault-fill veins curve outwards into the hangingwall of the fault zone to form sub-horizontal extension veins. The lack of substantial rotation of steeply-dipping bedding in the zone of vein curvature indicates that veins have grown with a curved geometry rather than being deformed into their curved shape. Field of view is about 4 meters wide.

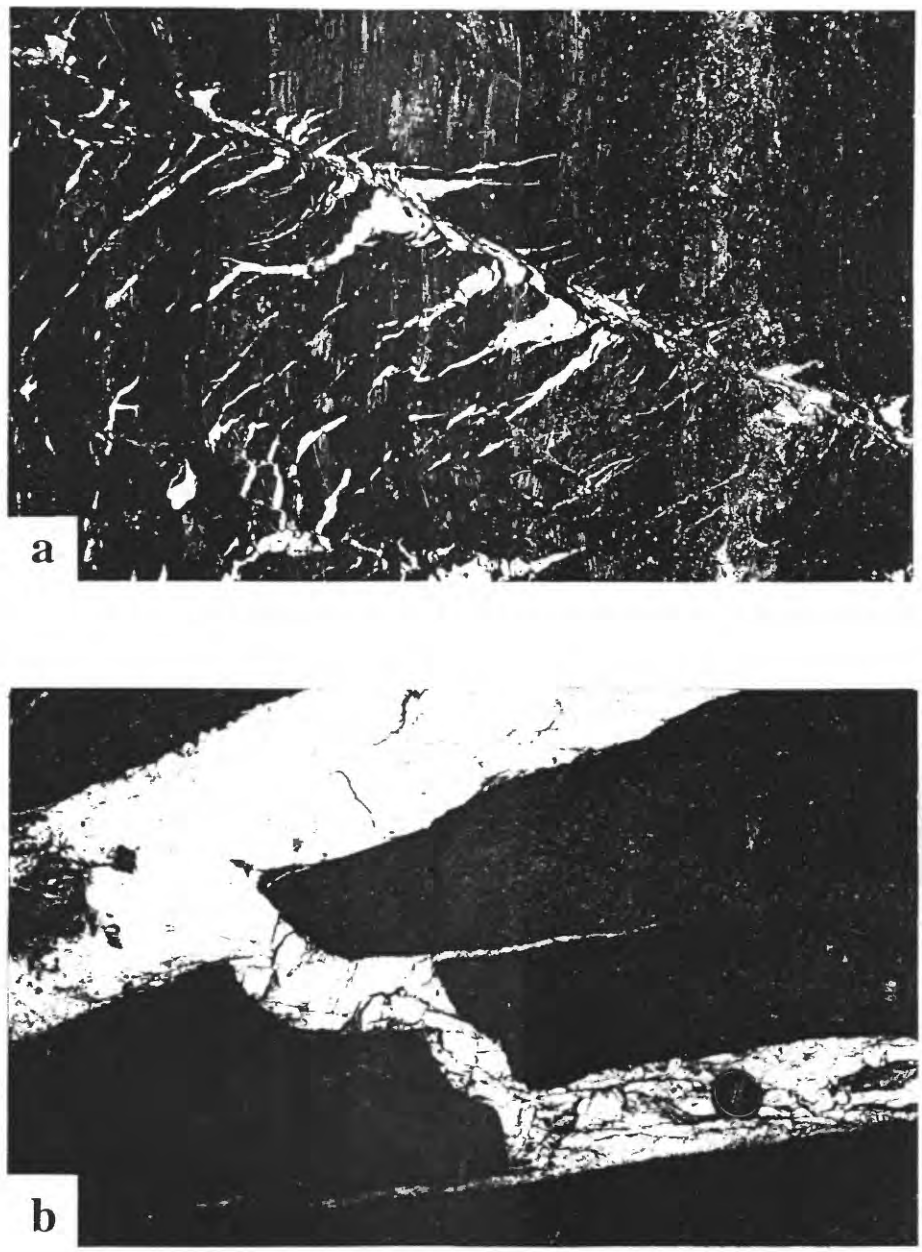



**Figure 3.**
**a.** Wedge-shaped extension veins propagating outwards from a small, bedding-discordant reverse fault. Field of view is about 1.5 meters wide.
**b.** Quartz extension veins in sandstone host-rock. Note the steeply-dipping vein segment linking two gently-dipping veins. Field of view is about 1.5 meters wide.

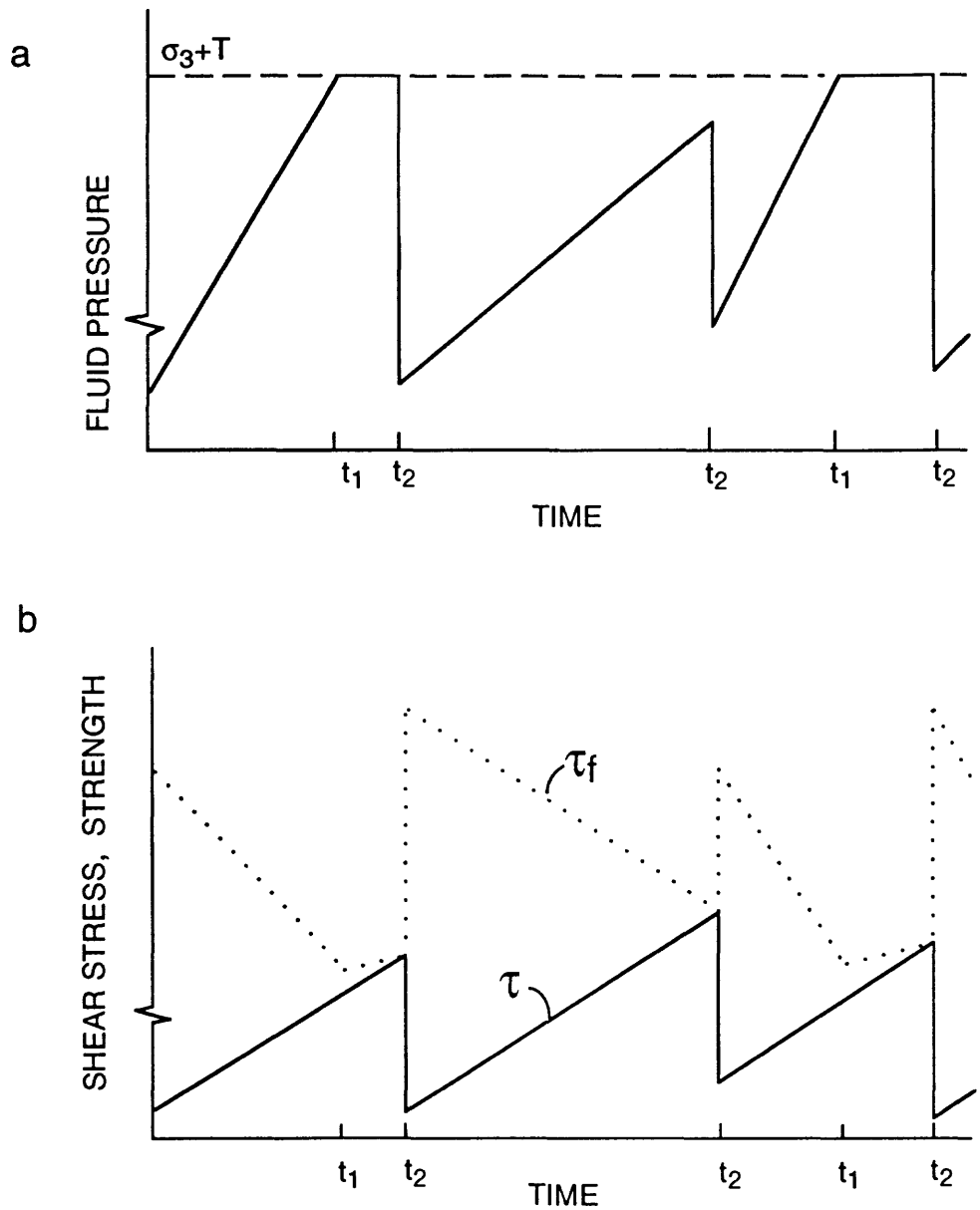



**Figure 4.** Illustration of cyclic variations in (**a**) fluid pressure, and (**b**) shear stress ($\tau$) and shear strength ($\tau_f$) as a function of time during fault-valve behavior in the WGFZ. Hydraulic extension fractures open adjacent to faults at times $t_1$ when fluid pressure equals $\sigma_3 + T$ ($\sigma_3$ is the minimum principal stress, T is the tensile strength of the intact wall-rock). Shear failure occurs at times $t_2$. Note the strong dependence of fault shear strength on both fluid pressure and shear stress. The fluid pressure history is influenced by several factors, including the rate of hydrothermal re-sealing and rate of fluid recharge within the fault after each failure event.

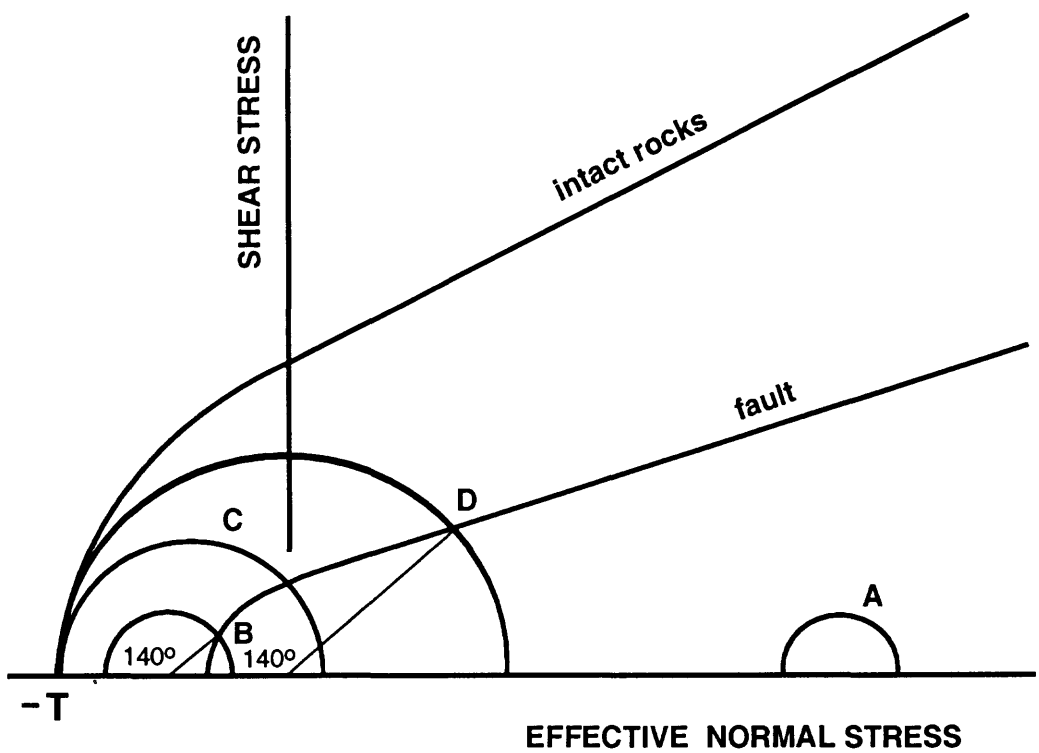



**Figure 5**. Mohr diagram schematically illustrating evolution of effective stress states during a fault-valve cycle in the WGFZ. Failure envelopes for both the fault and the intact wall-rock are shown. The fault is inclined at $70^{o}$ to $\sigma_1$. Immediately after a shear failure event, shear stress and fluid pressure are low (circle A). Localised post-seismic, hydrothermal sealing of the fault zone then progressively increases fluid pressures and decreases effective stresses at the same time as shear stresses begin to recover. Dilatation of steeply-dipping, bedding-parallel fault segments can occur at B. Opening of hydraulic fractures at supralithostatic fluid pressures occurs adjacent to the fault zone at C. With increasing shear stress, dilatant fractures close and shear failure of the fault zone occurs at D.

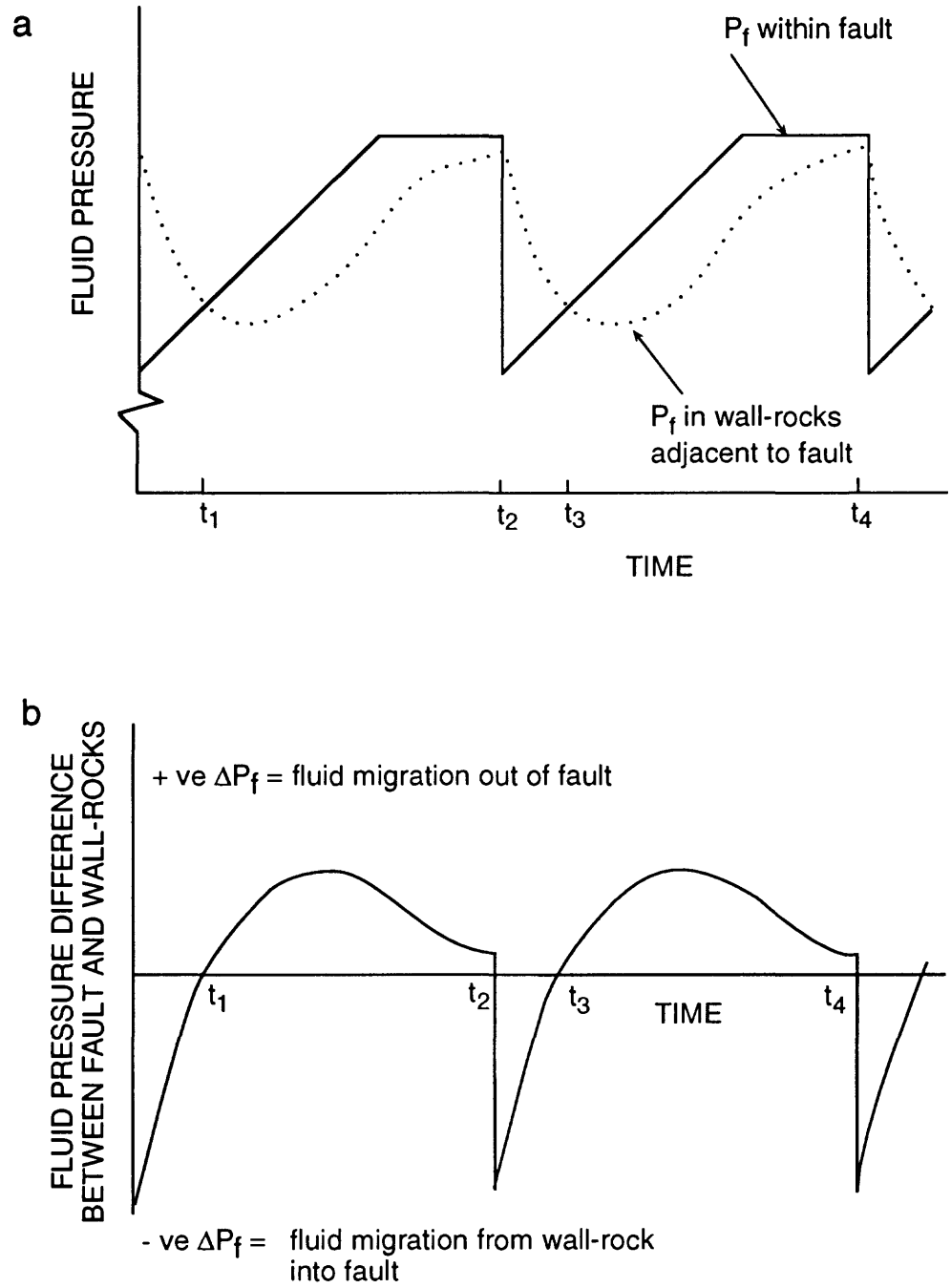



**Figure 6.**
**a.** Time-dependence of fluid pressure within a fault and in the adjacent wall-rocks . Slip episodes occur at times $t_2$ and $t_4$.
**b.** Time variation of hydraulic head adjacent to a fault. As pressure seals form prior to failure, fluid pressure in the fault zone becomes higher than in the adjacent wall-rocks at time $t_1$ and $t_3$. This promotes fluid migration out of the fault and into the wall-rocks. Valve breaching and opening of dilatant jogs during failure at times $t_2$ and $t_4$ cause fluid pressures in the fault to drop transiently below fluid pressure levels in the nearby wall-rocks. This causes fluids to migrate back into the fault zone where they can mix with fluids migrating up through the fault from deeper levels.

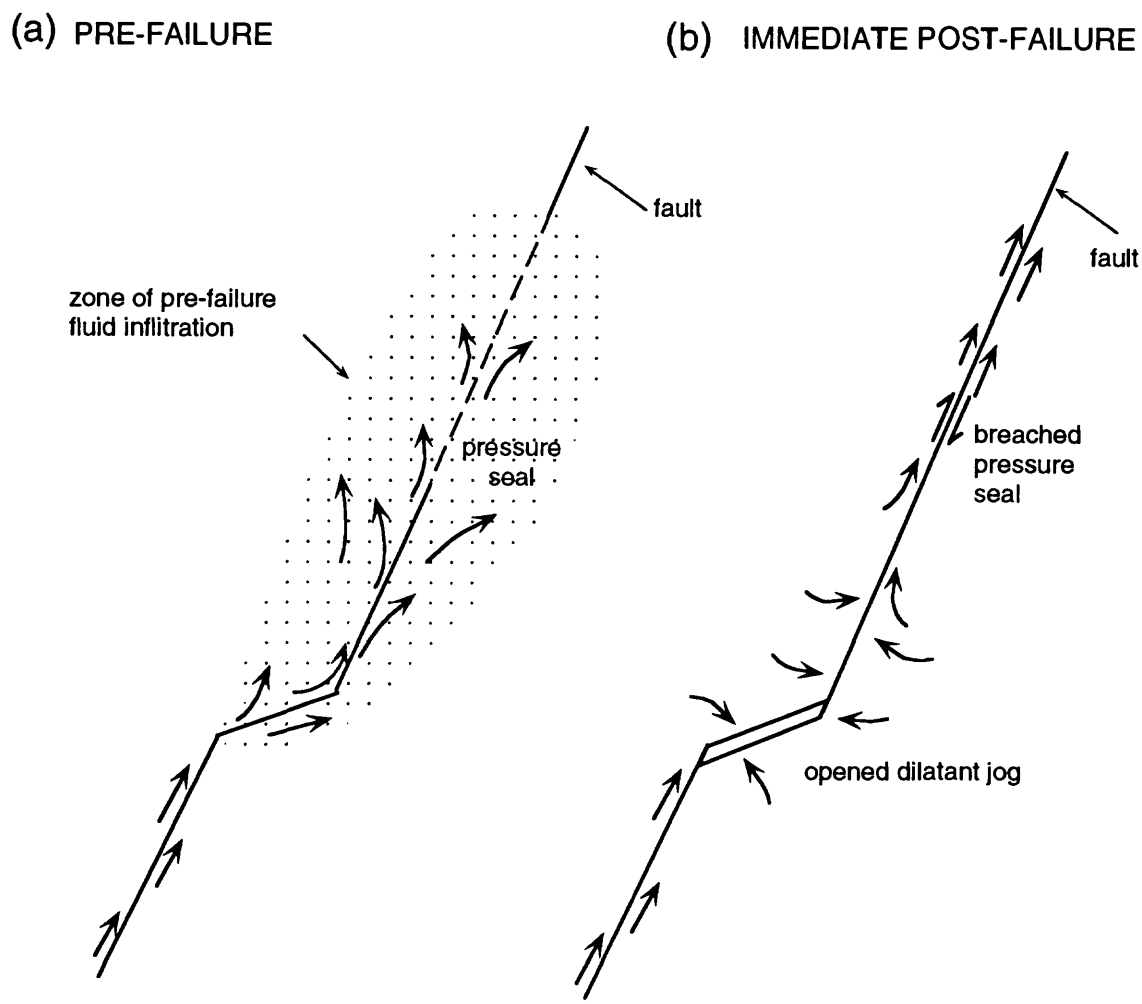



**Figure 7.** Schematic illustration of fluid migration patterns associated with cyclic fluid pressure fluctuations during fault-valve behavior.
**a.** Fluid migration out of the fault beneath a pressure seal prior to a slip event.
**b.** Fluid migration from wall-rocks back into the fault during the immediate post-rupture period.

# MESOTHERMAL GOLD-QUARTZ VEINS AND EARTHQUAKES

François Robert[1] and Anne-Marie Boullier[2]

1: Geological Survey of Canada, 601 Booth St., Ottawa, CANADA, K1A 0E8
2: C.R.P.G., C.N.R.S., B.P. 20, 54501 Vandoeuvre-les-Nancy, Cedex, FRANCE

## ABSTRACT

Mesothermal gold-quartz vein fields in greenstone terranes provide a record of fluid involvement during faulting and may be products of seismic processes near the base of the seismogenic regime. Detailed structural studies of these veins at Val d'Or in the Abitibi greenstone belt indicate lithostatic fluid pressures and cyclic stress reversals, and support the analogy between vein development and the earthquake stress cycle. A translation of selected gold-quartz vein characteristics into earthquake rupture parameters, combined with vein distribution and dimensions of the field they define, suggests that cyclic vein development was related to aftershock activity along subsidiary structures associated with major earthquake ruptures along an adjacent crustal-scale fault.

## INTRODUCTION

Mesothermal gold-quartz vein systems provide good evidence for fluid involvement during faulting. Their formation is interpreted to involve coupled stress- and fluid pressure-cycling, which have been correlated with the earthquake stress cycle (Sibson *et al.*, 1988; Cox *et al.*, 1990; Boullier and Robert, 1992). These veins may thus be considered as exhumed fossil products of seismic activity and offer the possibility of studying the effects and importance of fluids along seismically active faults, a subject of considerable interest (see Rice, 1992; Byerlee, 1993). Considered at the district-scale, these systems may also provide insight to larger scale fluid circulation and redistribution in the crust accompanying large earthquakes. The main objectives of this paper are to further explore these potential links between vein formation and earthquake processes and their implications, and to point out areas where further studies are needed.

Gold-quartz-tourmaline-carbonate (QTC) veins of the Val d'Or district in southeastern Abitibi greenstone belt, Canada, have been structurally documented in detail and offer a good opportunity for studying fault-related processes. The veins are well exposed in three dimensions through numerous mines and they typify the structural character of many other gold-quartz vein deposits and districts around the world (Eisenlohr *et al.*, 1989; Hodgson, 1989; Poulsen and Robert, 1989). The structural attributes of these gold-quartz veins and the dynamics of their development are first presented and serve as a basis for discussion their relations with earthquake processes.

## STRUCTURAL SETTING OF GOLD-QUARTZ VEINS

Like many other gold-quartz vein districts, Val d'Or is localized along a crustal-scale, first-order fault zone, the Larder Lake-Cadillac Fault (LLCF; Fig. 1). Such faults typically have a long history of movement and are considered to represent the main channelways for upward migration of deep fluids (Eisenlohr *et al.*, 1989; Kerrich and Wyman, 1990). Within districts, however, gold-quartz veins are rather associated with subsidiairy shear zones of second and third orders that are developed away from the first order fault. At Val d'Or, structural and relationships indicate that shear zones of all orders were formed (third order) or reactivated (first and second orders) during a late increment of deformation (Robert, 1990). The presence of similar veins and gold-related hydrothermal alteration in shear zones of all orders further indicates that they formed a district-scale, *3-D network* of interconnected shear zones as depicted in Figure 2.

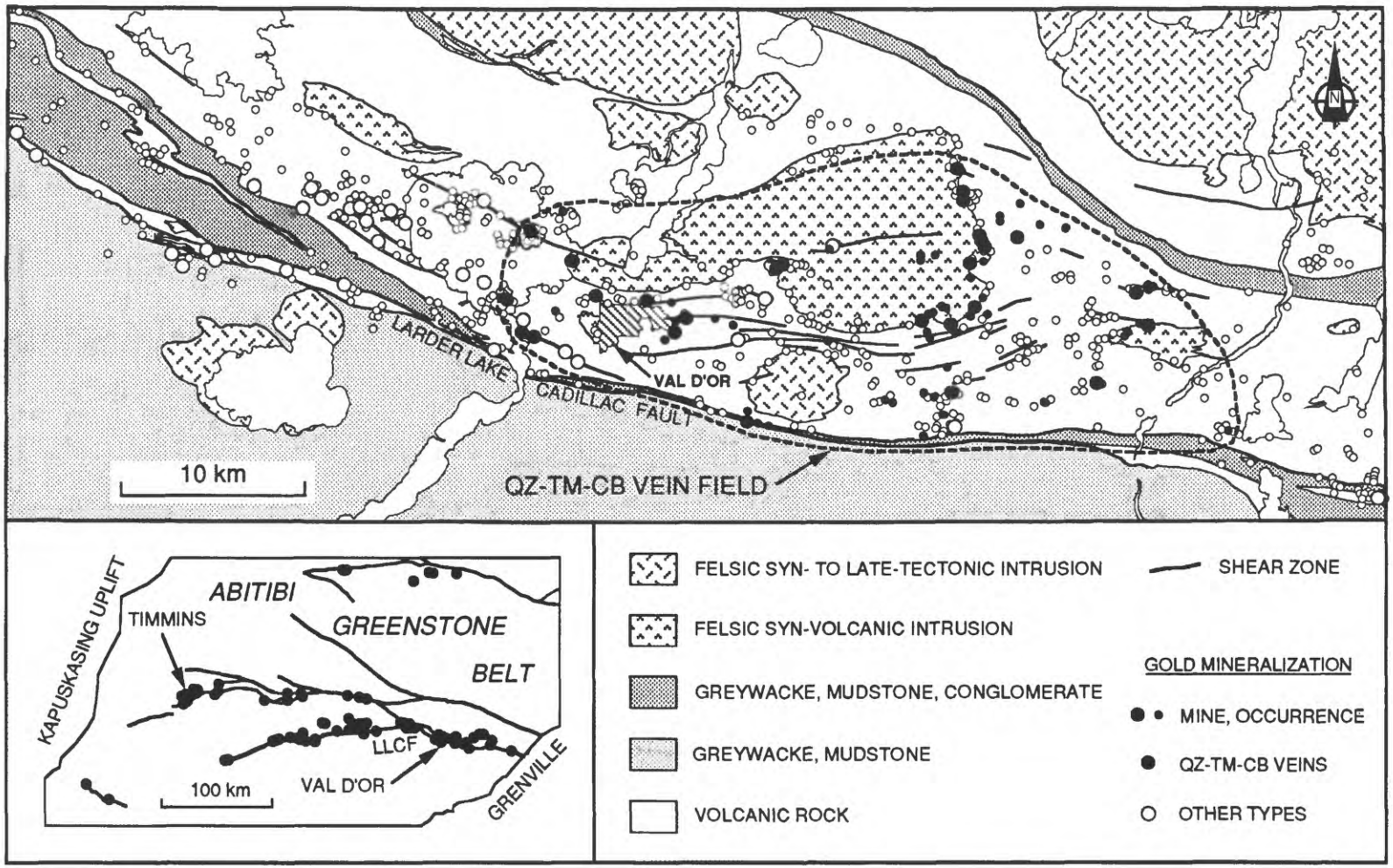



*Figure 1: Simplified geology of the Val d'Or area showing the distribution of gold mineralization. QZ-TM-CB veins occur in a restricted area and define an auriferous vein field. The inset shows the distribution of major fault zones and of significant gold deposits in the Abitibi greenstone belt.*

All shear zones have similar east-west strikes (Fig. 1) but they differ in dips: first- and second-order shear zones are subvertical, whereas third-order ones have dips in the range of 35-75° to the north or to the south (Fig. 2). All documented third-order shear zones are reverse to reverse-oblique and available data on first- and second-order shear zones suggest similar movements (Robert, 1990, 1993). The geometry and kinematics of these shear zones record northerly-directed bulk horizontal shortening (Fig. 2), compatible with the regional strain axes (Robert, 1990) and indicating development in a *compressional* regime. The only significant superimposed deformation in the district is a subsequent reactivation of the LLCF as a dextral strike-slip fault, which was accompanied by slight reactivation of a few third-order shear zones and veins (Robert, 1990).

Shear zones of all orders contain rocks that display mixed continuous-discontinuous mechanical behavior, compatible with the middle greenschist grade of the district. Such fault behavior and metamorphic grade are those expected near the base of the seismogenic regime in actively deforming crust (Sibson *et al.*, 1988).

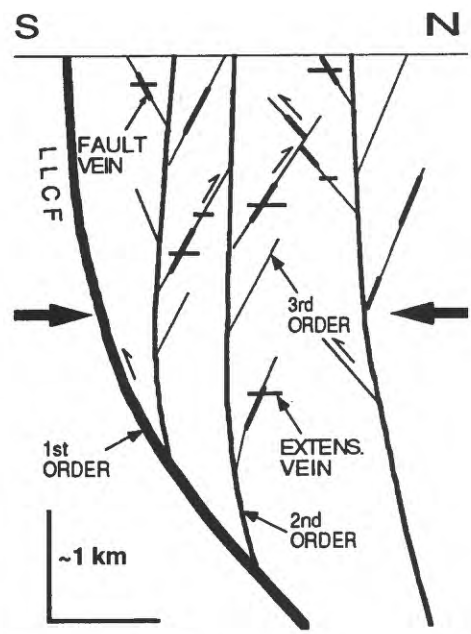



*Figure 2: Schematic cross-section through the Val d'Or district showing relationships between veins and different orders of shear zones.*

At Val d'Or, two main generations of shear zone-related gold-quartz veins are recognized. Early quartz-carbonate-pyrite veins occur chiefly in the western part of the district where they are associated with second-order shear zones. They are commonly folded and boudinaged and are cut by diorite and tonalite dikes. The younger quartz-tourmaline-carbonate-pyrite (QTC) veins considered here are typically associated with third-order shear zones which cut all intrusive rocks and are only overprinted by weak effects of the dextral strike-slip reactivation of the LLCF. All these QTC veins possess identical hydrothermal and structural characteristics. They occur within the area outlined in Figure 1 and define a *vein field*, developed during a late increment of regional shortening.

Such a vein field can be regarded as representing the extent of fluid "dispersion" around a zone of upward fluid migration along the adjacent first-order structure (LLCF). As determined from diamond drilling and from deep seismic reflection profiling (cf. Green *et al.*, 1990), the LLCF has a steep northerly dip (Fig. 2). The distribution of QTC veins only on the north side of the LLCF in the Val d'Or district (Fig. 1) is consistent with fluid leakage into the hangingwall of a main conduit into subsidiary structures.

## STRUCTURAL CHARACTERISTICS OF VEINS

Two main types of gold-quartz veins are particularly well developed among the QTC veins at Val d'Or: shear zone-hosted *fault veins* and fringing *extensional veins*. The gold deposits are networks of single to multiple sets of fault veins combined with extensional veins. As the structural characteristics of the QTC veins at Val d'Or have been described in detail by Robert and Brown (1986 a, b), Robert (1990), and Boullier and Robert (1992), only the most significant structural features are presented below.

## Mesostructures

*Fault veins* form elliptical lenses (25-100 m long x 1-2 m thick) within more continuous host shear zones or discrete faults (50-1000 m long) and have moderate to steep dips. They are typically laminated, reflecting incremental development, and consist of individual quartz laminae, (1-20 m long x <10 cm thick) separated either by discrete slip surfaces (Figs. 3A and B) or by slivers of foliated wallrocks of variable thicknesses (see Robert and Brown, 1986a and Robert, 1990). Striations on slip surfaces are common and are consistent with the slip directions determined along the host shear zones. Slickenfibres of quartz and tourmaline on slip surfaces further indicate reverse to reverse-oblique movement, again consistent with that along the host shear zone. These hydrothermal slickenfibres demonstrate that development of fault veins was accompanied by *active reverse slip*. The total amount of reverse slip along fault veins and their host structures could be determined in a small number of cases; it ranges from 0.1-2.0 m for structures on the order of 50-100 m of dip length, to 20-30 m along structures 500-1000 m of dip length (Robert, unpublished data). Such figures are probably maximum values for the amount of slip related to vein formation, because they include slip along the host shear zone or fault prior to vein development. Hydrothermal slickenfibres indicating *normal* movement are also observed on a small number of fault veins.

In addition, two types of breccia locally occur along or within fault veins. The first is a jigsaw puzzle or implosion breccia, which consists of unrotated angular wallrock fragments set in a hydrothermal matrix of fine grained massive tourmaline (Fig. 3C). The second can be regarded as a fault breccia: it forms narrow lenses within some fault veins and consists of angular but rotated fragments of veins and altered wallrocks in a hydrothermal matrix.

*Extensional veins* are sub-horizontal tabular bodies extending in intact rocks up to 75 m away from faults. They have high aspect ratios, covering areas of $10^2$-$10^4$ $m^2$ with thicknesses ranging from 1 cm to 1 m. Extensional veins generally consist of several wall-parallel ribbons of varying mineral proportions and textures (Fig. 3D), representing growth layers and attesting to their incremental development. Individual ribbons commonly display well developed open-space filling textures or mineral fibres (tourmaline +/- carbonate; Fig. 3D) and crack-seal; they reach thicknesses of 10 cm and areas of 1000 $m^2$. Lack of offsets of markers other than those that can be accomodated by pure dilation, the orientation of mineral fibres perpendicular to vein walls (Fig. 3D), and the common presence of rock bridges show that these veins are truly extensional.

Sigmoidal extensional veins of restricted lateral extent are also locally developed near the lateral and up- and down-dip terminations of fault veins, as illustrated in Figure 3B. These veins are interpreted to represent Mode I fractures developed at fault tips (see Pollard and Segall, 1987; Scholz, 1989); their common position at fault vein terminations and their sigmoidal shapes are consistent with that expected along reverse faults. Such sigmoidal extensional veins are distinguished from the throughgoing tabular extensional veins described above: tabular extensional veins, where not offset by subsequent reverse slip, can extend *on both sides* of a single fault vein and pinch away from it, in contrast to sigmoidal extensional veins which are typically restricted to one side of the fault (Fig. 3C). Given their orientation and formation in a compressional stress field, tabular extensional veins are best interpreted as hydraulic fractures (see discussion below).

Mutually cross-cuting relationships are observed between fault veins and extensional veins, commonly along the same fault vein and locally at a single site (see Robert and Brown, 1986a). Such relationships indicate that the two vein types are broadly contemporaneous and developed in a cyclic, sequential manner (Robert, 1990; Boullier and Robert 1992). The presence of hydrothermal slickenfibres and sigmoidal extensional veins at fault vein terminations further indicate that vein development accompanied active faulting.

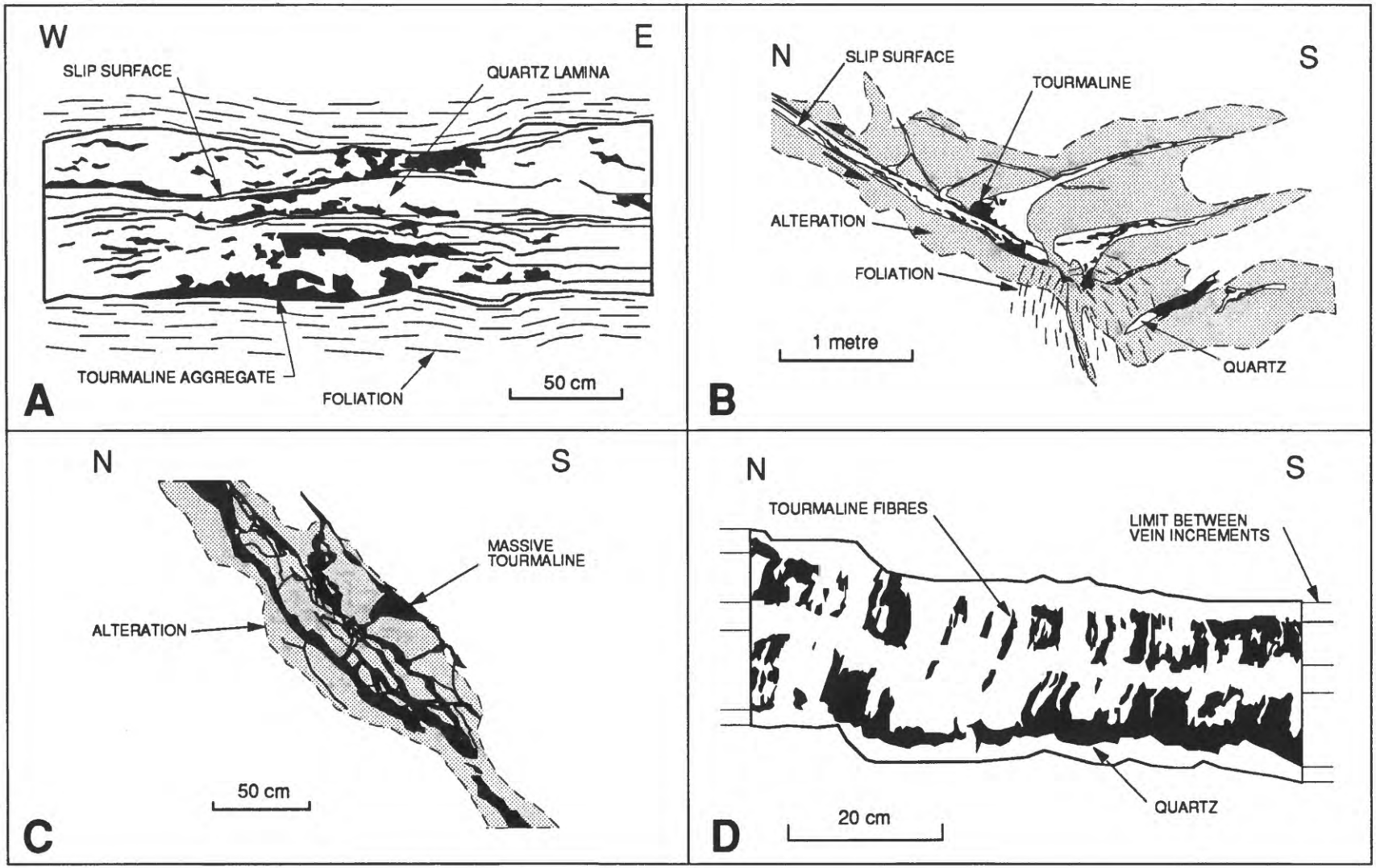



*Figure 3: Examples of QTC veins, traced from photographs. A- Horizontal plan view of a fault vein with abundant slip surfaces; the "clean" quartz lamina in the center of the vein represents a late increment of opening. B- Vertical cross-section of down-dip termination of a reverse fault vein with related sigmoidal extensional veins. Truncated extensional veins likely formed during an earlier increment of fault vein propagation. C- Vertical cross-section of a jigsaw puzzle breccia. D- Vertical cross-section of a subhorizontal extensional vein with well developed vertical tourmaline fibres; note the multiple opening increments.*

## Microstructures

Microstructures of both types of veins have been examined in detail by Boullier and Robert (1992) and reflect the interplay of stages of vein growth and vein deformation (Table 1). Growth microstructures are best preserved in tabular extensional veins extending outside shear zones, whereas fault veins are dominated by deformation microstructures, consistent with their development in active faults.

In extensional veins, growth microstructures include open-space filling, mineral fibres and crack-seal, the latter two recording *vertical dilation* corresponding to opening of the veins (Table 1). Individual ribbons or segments of ribbons are dominated by one of these three growth structures and it is postulated that the type of microstructure produced reflects differences between the rate of wall separation and that of mineral precipitation, as proposed by Cox (1991) and Foxford *et al.* (1991). In addition to undulose extinction and prismatic subgrains in quartz, deformation microstructures include kinks developed in mineral fibres and crystals at high angle to vein walls, as well as stretching of crystals parallel to vein walls (see Boullier and Robert, 1992). Such deformation microstructures reflect *vertical shortening* of the extensional veins (see below).

In fault veins, rare open-space filling textures and crack seal structures are the only observed growth features (Table 1). Most quartz laminae are dominated by prismatic subgrains or completely polygonized crystals. Stylolites are developed along the walls of laminae; the north-south, horizontal orientation of their peaks is consistent with the maximum principal stress direction at the time of vein development (Fig. 2). Stylolites indicate aseismic (low strain rate) deformation attributed to horizontal shortening and vertical extension, whereas slip planes on very fine grained crushed tourmaline layers indicate seismic (high strain rate) deformation (Gratier and Gamond, 1989). Microstructures produced by such contrasted processes are juxtaposed within single thin sections.

*Table 1: summary of vein microstructures and their significance.*

| | FAULT VEINS | | EXTENSIONAL VEINS | |
|---|---|---|---|---|
| GROWTH STAGE | open-space filling | | open-space filling | |
| | crack seal (rare) | | crack-seal | *vertical* |
| | | | mineral fibres | *dilation* |
| | | | horizontal F.I. planes | |
| DEFORMATION STAGE | stylolites | | kinked fibres | *vertical* |
| | plastic deformation | *aseismic* | vertical F.I. planes | *shortening* |
| | dynamic recryst. | | horizontal boudinage | |
| | F.I. planes | | | |
| | slip planes | *seismic* | | |
| | crushed tourmaline | | | |

Healed microcracks, outlined by planes of fluid inclusions occur with systematic orientations in both vein types. Such microcracks have been shown to be of Mode I (Boullier and Robert, 1992) and their different orientations in the veins must reflect different orientations of the local stress field at different times. In extensional veins, healed microcracks occur both parallel and perpendicular to the vein walls (Fig. 4A), indicating *vertical dilation* and *vertical shortening*, respectively. In fault veins, a large number of microcracks are perpendicular to, and record stretching along, the slip direction (Fig. 4B).

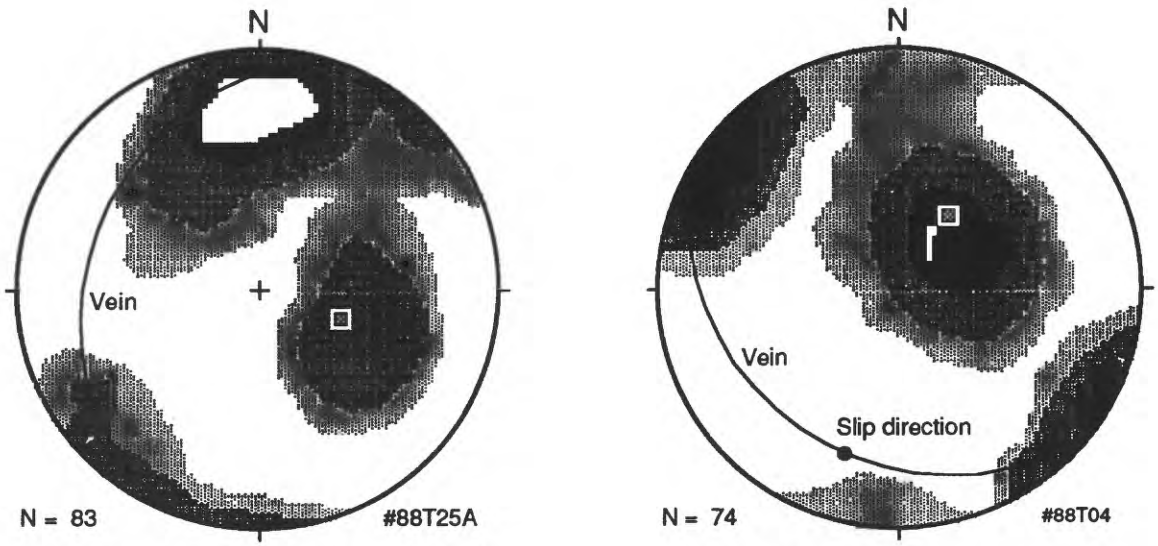



*Figure 4: Stereographic projections (equal area) of poles of secondary fluid inclusion planes; Perron mine. Kamb contours. In both types of veins, planes are oriented parallel and perpendicular to the vein walls.*

Significant variations in degree of plastic deformation and in abundance of microcracks are observed between adjacent ribbons in extensional veins and between adjacent quartz laminae in fault veins. Truncations of "old" microcracks by newly added material are also observed in both vein types. These observations indicate that stages of vein growth are alternating with stages of vein deformation and that both stages are part of a repeating cycle leading to the development of ribboned or laminated veins (Boullier and Robert, 1992). They also verify that the observed deformation of vein materials is not the result of superimposed post-vein strain.

## INTERPRETATION OF CYCLIC VEIN DEVELOPMENT

The structural attributes and relationships presented above clearly point to the sequential, cyclic development of reverse fault veins and extensional veins in alternating stages of growth and deformation. Reverse slip along fault veins and growth of extensional veins (hydraulic fractures) require a compressional regime with horizontal $\sigma 1$ and vertical $\sigma 3$, consistent with regional strain patterns (Robert, 1990). It further requires conditions of near-lithostatic Pf and low differential stress ($\sigma 1 - \sigma 3 < 4T$ and $Pf > \sigma 3 + T$). However, the documented vertical shortening of extensional veins and hydrothermal slickenfibres of normal sense along a few fault veins also indicate periods of vertical $\sigma 1$ and horizontal $\sigma 3$, which require at least transient and local stress reversals during vein development. Vein formation has been correlated with the earthquake cycle in the *fault-valve* model (Sibson *et al.*, 1988; Cox *et al.*, 1991; Boullier and Robert, 1992) and is here regarded as a two stage process in which mesoscopic and microscopic vein structures are integrated (Figure 5).

Considering a fault vein sealed by hydrothermal precipitation following rupture, the *interseismic period* of the cycle (after earthquake-related Pf transients have ceased) is characterized by restoring $\tau$ and the pressure of externally derived fluid along the fault. Extensional veins will grow adjacent to faults at this stage (Fig. 5A), once Pf reaches supra-lithostatic levels. However, the development of these veins will be influenced by the relative rates of Pf and $\tau$ restoration, the

degree of misorientation (dip) of the fault and its cohesive strength (Cox, 1991; Sibson, 1992). For example, if $\tau$ along the fault restores faster than Pf, then rupture may occur along the fault (depending on its degree of misorientation) before development of extensional veins. This may explain why only fault veins are observed in some deposits.

Two different types of microcracks are preserved in extensional veins: delicate crack-seal layers and healed microcracks. The first type has been interpreted to result from sub-critical crack growth by Cox (1991) and by Boullier and Robert (1992); it is more probable that the second results from dynamic crack growth. The difference in propagation mode may reflect differences in driving pressure (Pollard and Segall, 1987; Cox, 1991) and differences in the degree of fluid overpressure between cycles. Within fault veins, the interseismic period is marked by plastic deformation of existing quartz laminae and stylolitization related to increase in $\tau$. The development of reverse hydrothermal slickenfibres may record aseismic slip during that period.

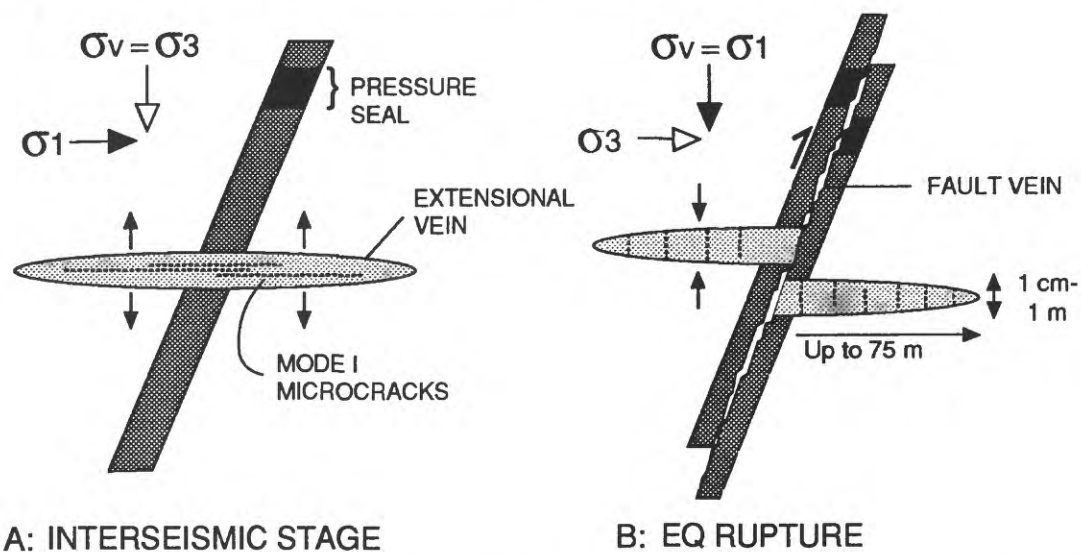



*Figure 5: Diagram illustrating the main stages of vein development and selected microstructures in relation to the earthquake cycle. See text for explanations.*

*Rupture* along the fault will occur once $\tau$ reaches the appropriate level given the shear strength of the fault and the prevailing Pf. An important consequence of rupture is the breaching of the pressure seal, drainage of overpressured fluid and drop of Pf. Rupture will also be accompanied by release of $\tau$ along the fault and by the transient reversal of local principal stress axes as documented above (Fig. 5B; Boullier and Robert, 1992), which is conceivable under conditions of low differential stress. In turn, such stress reversals require *total release* of $\tau$ along the fault plane and implies dynamic overshoot (Sibson, 1992). The formation of hydrothermal slickenfibres of normal sense observed along some fault veins may be a consequence of overshoot.

The post-rupture Pf drop will likely promote precipitation of quartz and carbonate, whose solubilities are sentitive to pressure under those hydrothermal conditions, and lead to progressive resealing of the fault (Sibson et al., 1988). If Pf restores much faster than $\tau$ along the fault, the local environment can be considered to be under a nearly isotropic stress field, and it is conceivable that dilation of weak surfaces could take place, including foliation planes in the host shear zones or the fault vein itself, irrespective of their orientation.

According to this model, each recognizable quartz laminae within a fault vein should be the product of one cycle. Within fault veins, several tens (20-100) of laminae can typically be identified. They provide a *minimum* estimate for the number of cycles, or earthquakes, involved in the formation of these veins, because individual quartz laminae not bounded by tourmaline-bearing surfaces will be very difficult to recognize (quartz against quartz) and because not all slip increments are likely to be accompanied by dilation and precipitation of quartz.

## DISCUSSION

The two-stage fault valve model presented above hinges heavily on two points which deserve further consideration: lithostatic fluid pressures and seismic slip events within fault veins.

The best evidence for lithostatic fluid pressures comes from the presence of subhorizontal tabular extensional veins extending in intact rocks on both sides of fault veins. As argued by a number of authors, tensile fracturing at depths corresponding to greenschist metamorphic conditions require fluid pressures exceeding the minimum principal compressive stress (see for example Secor, 1969; Etheridge *et al.*, 1984). In a large number of ribbons within horizontal extensional veins at Val d'Or, the presence and preservation of delicate open-space filling textures, such as rosettes of fine tourmaline crystals up to 10 cm long, in horizontal extensional veins indicates minimum wall separation of several centimetres and that the fracture remained opened until complete filling (Robert and Brown, 1986b). The dimensions of such open horizontal fractures, given by those of the corresponding ribbons, are several centimetres thick and reach areas of $10^3$ $m^2$. This requires fluid pressures at least equal to the lithostatic load, i.e. sufficient to lift the weight of the overlying column of rocks.

Direct evidence of seismic slip within fault veins is scarce, perhaps because of a poor degree of preservation. For example, cataclasite developed within quartz could be easily overprinted during crystal plastic deformation related to subsequent shear stress accumulation, or by fluid-enhanced recrystallization and grain boundary migration during a stage of stress relaxation. However, discrete slip surfaces and sub-millimetric bands of finely crushed (cataclastic?) tourmaline within fault veins may record seismic slip. Perhaps the best evidence comes from the fault breccias with hydrothermal cement that occur within or along a small number of fault veins. Finally, the presence of implosion breccias such as that illustrated in Figure 3C, as well as the documented stress reversal during vein development are best explained in terms of a *sudden* release of shear stress following an EQ and the consequent Pf drop. Perhaps one of the most convincing tests of the fault valve model would be to document differences in trapping pressures (reflecting Pf fluctuations) of fluid inclusions between pairs of extensional and fault veins. Studies to that effect are presently under way.

Accepting the proposed fault valve model, one can further explore links between vein formation and EQ processes and speculate about their implications, *at least in a preliminary way.*

### Magnitude of vein-related earthquakes

In considering the analogy with EQ processes, one of the first questions that comes to mind is the magnitude of earthquakes accompanying each seismic slip increment along a fault vein. A crude estimate can be obtained based on the existing relationship between EQ magnitude and typical source parameters such as rupture area and amount of slip, as determined by Sibson (1989, his Table 1) for standard earthquakes with shear stress drops of ~30 bars. These parameters can be estimated for a small number of fault veins, assuming they were derived from such standard events.

The maximum sizes of rupture areas are constrained by the total dimensions of the third-order structures hosting fault veins, which fall in the range of $10^4$-$10^6$ $m^2$. This is a maximum because any rupture is unlikely to occupy the full dimensions of the finite fault. Third-order structures and

fault veins are low-displacement structures: based on a small number of adequately constrained cases, total offsets, including any pre-vein slip, range from 0.1 to 2 m along small, discrete faults to a maximum of 20-30 m along km-scale ductile shear zones. Fault veins hosted in discrete faults without ductile components of deformation provide the most realistic estimates of vein-related total slip. Assuming seismic slip only, the amount of slip related to each cycle is obtained by dividing the total displacements by the number of slip increments recognized within fault veins. As previously pointed out, the minimum number of slip increments along fault veins ranges from 20 to 100. Combined with intermediate values of total displacement along small, discrete faults, such number of slip increments yield maximum displacements of a few centimetres for each slip event.

Rupture parameters crudely estimated using the above constraints, i.e. a few centimetres of slip and rupture areas of $10^4$-$10^6$ $m^2$, correspond to earthquakes of M3 to M4 or less (cf. Table 1 in Sibson, 1989). Given that these rupture parameters represent maximum estimates, vein-related earthquakes are likely to have been of smaller magnitudes. Clearly, additional studies of well exposed fault veins should aim at documenting the number of slip increments, the total offset along the vein, and the dimensions of the host structure and those of individual quartz laminae.

## QTC vein distribution vs earthquake distribution

The magnitudes of vein-related earthquakes are relatively small compared to the large earthquakes expected along lithospheric fault zones such as LLCF near the base of seismogenic regime (Sibson, 1989). The relationships between expected large earthquakes along 1st-order structures and the widely distributed, smaller vein-related earthquakes along subsidiary structures require further investigation.

The distribution of QTC veins within the vein field at Val d'Or (Fig. 1) can be regarded as reflecting the distribution of small earthquake events, integrated during the lifetime of the hydrothermal system. The QTC veins at Val d'Or (Fig. 1) form a field or a cluster which is approximately 45 km long and 15 km across and which is located in the hangingwall of the LLCF (Fig. 1). Such dimensions compare very well with those of clusters of aftershocks following major ruptures along the San Andreas fault system, such as the Loma Prieta, California, earthquake of October 17, 1989, with a cluster approximately 40 km long and 5 km across (McNutt, 1990), or the Morgan Hill, California, earthquake of April 24, 1984, with a cluster approximately 40 km long and 13 km across (Cockerham and Eaton, 1987). It is also important to note that in both examples, at least some of the aftershocks occurred along subsidiary structures. The magnitude of the vein-related earthquake estimated above, M3-4 or less, is consistent with the magnitudes of the majority of aftershocks. This raises the interesting possibility that the individual seismic slip events along fault veins correspond to aftershocks along subsidiary structures related to major ruptures along the LLCF. As pointed out by Sibson (1993), changes in mean stress related to large earthquakes along lithospheric faults induce fluid redistribution around the faults in fluid saturated crust. Such effects may play a role in the district-scale dispersion of mineralizing fluids in the network of interconnected subsidiary structures and should be further explored.

## Duration of the cycle

Another immediate question related to the two-stage fault valve model presented here concerns the duration of the cycle, which would broadly correspond to the duration of the interseismic period.

Perhaps constraints could be placed from the inferred mode and rate of fracture propagation and from the dimensions of individual ribbons dominated by crack-seal textures. As indicated above, the delicate crack-seal structures must reflect slow, sub-critical crack growth. Individual crack-seal-dominated ribbons typically consist of 1000 increments of an average thickness of 25 microns, and can be traced laterally up to at least 10 m. We consider that development of the entire

ribbon takes place during a single interseismic period, rather than each crack seal increment representing one complete cycle. The constant thickness of crack-seal increments and the constant shape, mimicking irregularities of the crack walls, over 1000 cyclic repetitions suggest a steady-state regime of formation. The observed lateral transition from crack-seal to open-space filling internal structure in single ribbons further supports our interpretations: the existence of open spaces requires that the fracture remains open until complete filling without completely closing, hence without Pf drop and within a single cycle .

Preliminary calculations, assuming sub-critical fracture propagation, for a typical crack-seal layer with the parameters indicated above and using experimentally determined crack velocity ($10^{-3}$ to $10^{-6}$ $m.s^{-1}$, Atkinson and Meredith, 1987), indicate that a ribbon would take from tens of days to hundreds of years to develop. Such duration of the interseismic period overlaps with the recurrence interval of large earthquakes along lithospheric fault zones.

## CONCLUSIONS

There is good evidence that mesothermal gold-quartz vein fields are fossil products of cyclic and linked hydrothermal and seismic activity. A preliminary translation of vein-related parametres into earthquake rupture parameters reveals many similarities of scale, dimensions and processes. Such approach is worth pursuing with more quantitative documentation of the physical and structural aspects of the veins and their host structures. It warrants further systematic quantitative studies of vein parameters. Finally, it should be emphasized that these vein fields offer excellent opportunities for exploring and studying earthquake-related processes in three dimensions and at different scales.

## ACKNOWLEDGEMENTS

Many of the ideas presented here evolved from, or were clarified by, discussions with Bill Brisbin, Stephen Cox, Yves Guégen, Jean-Pierre Gratier, Howard Poulsen and Richard Sibson. We would also like thank Howard Poulsen and Rick Sibson for their critical review of the manuscript, and Bill Brisbin for pointing out the similarity of scale between the vein field at Val d'Or and clusters of aftershocks related to major earthquake ruptures. AMB ackowledges funding from INSU-CNRS. This is DBT, Instabilités, contribution #626, CRPG contribution #1009 and GSC contribution #29493.

## REFERENCES

Atkinson B.K., and Meredith P.G. 1987. The theory of sub-critical crack growth with applications to minerals and rocks. *In* Fracture Mechanics of Rocks, *edited by* B.K. Atkinson. Academic Press, London, 111-166.

Boullier A.-M. and Robert F. 1992. Paloeseismic events recorded in Archean gold-quartz vein networks, Val d'Or Abitibi, Quebec. Journal of Structural Geology **14**, p. 161-179.

Byerlee J. 1993. Model for episodic flow of high-pressure water in fault zones before earthquakes. Geology **21**, 303-306.

Cockerman R.S. and Eaton J.P. 1987. The earthquake and its aftershocks, April 24 through September 30, 1984. U.S. Geological Survey Bulletin **1639**, 15-28.

Cox S.F. 1991. Geometry and internal structures of mesothermal vein systems: implications for hydrodynamics and ore genesis during deformation. U. Western Australia Publication **25**, p. 47-53.

Cox S.F., Etheridge M.E. and Wall V.J. 1990. Fluid pressure regimes and fluid dynamics during deformation of low-grade metamorphic terranes: implications for the genesis of mesothermal gold deposits; NUNA Conference Volume, Geological Association of Canada, p. 46-53.

Cox S.F., Wall V.J., Etheridge M.E., and Potter T.F.1991. Deformational and metamorphic processes in the formation of mesothermal vein-hosted gold deposits - examples from the Lachland fold belt in Central Victoria, Australia. Ore Geology Reviews **6**, 391-423.

Eisenlohr B.N., Groves D.I., and Partington G.A. 1989. Crustal-scale shear zones and their significance to Archean gold mineralization in Western Australia. Mineralium Deposita **24**, 1-8.

Etheridge M.A., Wall, V.J., Cox, S.F., and Vernon R.H. 1984. High fluid pressures during regional metamorphism and deformation: Implications for mass transport and deformation mechanisms. Journal of Geophysical Research **89**, no. B6, 4344-4358.

Foxford K.A., Nicholson R., and Polya D.A. 1991. Textural evolution of W-Cu-Sn-bearing hydrothermal veins at Minas da Panasqueira, Portugal. Mineralogical Magazine **55**, 435-445.

Gratier J.-P., and Gamond J.F. 1989. Transition betwen seismic and aseismic deformation in the upper crust. *In*: Deformation Mechanisms, Rheology and Tectonics, *edited by* R.J. Knipe and E.H. Rutter. Geological Society of London, Special Publication **54**, 461-473.

Green A.G., Milkreit B., Mayrand L.J., Ludden J.N., Hubert C., Jackson S.L., Sutcliffe R.H., West G.F., Verpaelst P., and Simard A. 1990. Deep structure of an Archean greenstone terrane. Nature **344**, 327-330.

Hodgson C.J. 1989. The structure of shear-related, vein-type gold deposits: A review. Ore Geology Reviews **4**, 231-273.

Kerrich R. and Wyman D. 1990. Geodynamic setting of mesothermal gold deposits: An association with accretionary tectonic regimes. Geology **18**, 882-882.

McNutt S., 1990 Loma Prieta earthquake, October 17, 1989, Santa Cruz County, California. California Geology, January 1990, p. 3-7.

Pollard D.P. and Segall P. 1987. Theoretical displacements and stresses near fractures in rocks: With applications to faults, joints, veins, dikes, and solution surfaces. In Fracture Mechanics of Rocks, edited by B.K. Atkinson. Academic Press Inc. (London) Ltd., p. 277-349.

Poulsen K. H., and Robert F. 1989. Shear zones and gold: practical examples from the southern Canadian Shield. *In* Mineralisation and Shear Zones. *Edited by* J.T. Burnsnall. Geological Association of Canada, Short Course Notes 6, pp. 239-266.

Rice J.R. 1992. Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault. *In* Fault Mechanics and Transport Properties in Rocks, *edited by* B. Evans and T.-F. Wong. Academic Press, 475-503.

Robert F. 1990. Structural setting and control of gold-quartz veins of the Val d'Or area, southeastern Abitibi Subprovince. *U. Western Australia Publication* **24**, 164-209.

Robert F. 1993. The geometrical relationship between the stretching lineation and the movement direction of shear zones: Discussion. Journal of Structural Geology **15**, 239-240.

Robert F., and Brown A.C., 1986a. Archean gold-bearing quartz veins at the Sigma mine, Abitibi greenstone belt, Quebec. Part I. Geologic relations and formation of the vein system. Economic Geology 81, 578-592.

Robert F., and Brown A.C., 1986b. Archean gold-bearing quartz veins at the Sigma mine, Abitibi greenstone belt, Quebec. Part II. Vein paragenesis and hydrothermal alteration. Economic Geology 81, 593-616.

Scholz C.H. 1989. Mechanics of faulting. Annual Review of Earth and Planetary Sciences **7**, 309-334.

Secor D.T. 1969. Mechanics of natural extension fracturing at depth in the Earth's crust. Geological Survey of Canada, Paper 68-52, 3-48.

Sibson R.H. 1989. Earthquake faulting as a structural process. Journal of Structural Geology **11**, 1-14.

Sibson R.H. 1992. Implications of fault-valve behavior for rupture nucleation and recurrence. Tectonophysics **211**, 283-293.
Sibson R.H. 1993. Load-strengthening versus load-weakening faulting. Journal of Structural Geology **15**, 123-128.
Sibson R.H., Robert F. and Poulsen K.H. 1988. High angle reverse faults, fluid-pressure cycling, and mesothermal gold-quartz deposits. Geology **16**, 551-555.

# Geological evidence for fluid involvement in the rupture processes of crustal earthquakes

Richard H. Sibson

*Department of Geology, University of Otago*
*P.O. Box 56, Dunedin, New Zealand*

**Abstract:** Evidence from shear zone mineral assemblages and hydrothermal veining suggests that major fault zones act as conduits for the passage of large fluid volumes at all crustal levels. "Dry" conditions for faulting are comparatively rare and are most likely to prevail at the time of fault zone inception in crystalline crust. Most active crustal faults are therefore expected to operate with fluid pressures at least at hydrostatic levels in their upper few kilometres, but there is geological evidence that suprahydrostatic (and, locally, lithostatic) fluid pressures operate in the lower levels of the seismogenic zone, and that *fault-valve* action is widespread. Incremental depositional textures in veins suggest that flow is predominantly episodic at seismogenic depths and that a dynamic interplay exists between earthquake faulting and fluid redistribution in the earth's crust. A range of dilatancy pumping mechanisms, variously dependent on $\Delta\tau$, $\Delta\bar{\sigma}$, and $\Delta P_f$, may contribute to cyclic fluid redistribution around faults but their scale of operation and relative effectiveness at different crustal levels remains unclear. However, it is clear that fault and fracture permeability must continually be renewed for them to remain effective channelways, and that stress cycling acts to *modulate* fluid flow in deforming upper crust affecting flow systems driven by long-term hydraulic potentials.

There is geological evidence for mechanical involvement of fluids at all stages of the earthquake cycle. Transient fluid pressure reductions at dilational jogs and bends may contribute to rupture arrest, and stress-controlled fluid redistribution influences the time-dependence of aftershock and swarm activity. The competition between creation and destruction of fracture permeability plays a critical role in the accumulation of fluid overpressures in the lower half of the seismogenic zone, affecting earthquake nucleation and recurrence. As a consequence of fluid pressure cycling from *fault-valve* action, rheological models of fault zones and crustal shear strength profiles must also be considered time-dependent, with integrated shear strength at a minimum prefailure but attaining a maximum value postfailure at the end of the discharge phase.

## Fluid pressure levels and fault instability

Frictional shear strength of an existing fault may be represented by a criterion of Coulomb form:

$$\tau_f = C + \mu_s\sigma_n' = C + \mu_s(\sigma_n - P_f) \qquad (1)$$

where C is the cohesive or cementation strength of the fault, $\mu_s$ is the static coefficient of friction (typically ~ 0.75), and $\sigma_n$ is the normal stress on the fault. Fault reactivation may thus be induced by rising shear stress, decreasing normal stress, or by increasing fluid pressure. From the time of Hubbert & Rubey's (1959) seminal paper on the mechanics of low-angle thrusting, it has been clear that overpressured (suprahydrostatic) fluids play a critical role in faulting and there is now good evidence that seismic rupturing is sometimes occurring within overpressured portions of the

crust such as the Western Taiwan fold/thrust belt and regions of California adjacent to the San Andreas fault (Sibson, 1990).

Processes of permeability destruction - diffusional crack-healing (Smith & Evans, 1984), porosity destruction through solution-precipitation processes (Sprunt & Nur, 1977; Morrow *et al.*, 1981), and hydrothermal self-sealing of fractures (Grindley & Browne, 1976; Batzle & Simmons, 1977) - are fast-acting at $T > 200°C$ in the lower half of the continental seismogenic zone, allowing for complex coupling between episodes of fault failure, the creation and destruction of fracture permeability, and fluid redistribution (e.g. Chester *et al.* 1993) (Fig. 1). In such regions, *fault-valve* action may be expected to occur when ruptures transect suprahydrostatic gradients in fluid pressure and breach permeability barriers, leading to upwards fluid discharge along the transient permeability of the fault zone and local reversion towards a hydrostatic fluid pressure gradient (Fig. 4). Fault-valve behavior thus leads to fluctuations in fault zone permeability, fluid pressure and strength tied to the earthquake stress cycle. Evidence for extensive fluid pressure cycling near fault zones comes from fluid inclusion studies (Mullis, 1988; Grant *et al.*, 1990; Parry & Bruhn, 1990; Boullier & Robert, 1992), and from local variations in very-low-grade metamorphic assemblages associated with brittle fracturing in the upper crust (Coombs, 1993).

**Possible dilatancy effects related to the fault loading cycle**

The state of dilatational strain in a rock mass may be affected by variations in the levels of both shear stress and effective mean stress, $\bar{\sigma}' = (\bar{\sigma} - P_f)$. Different dilatancy mechanisms may be grouped into those sensitive primarily to varying shear stress, $\Delta\tau$, those sensitive to variations in mean stress, $\Delta\bar{\sigma}$, and those that are driven by fluctuating fluid pressures, $\Delta P_f$ (Table 1). At present, the relative contributions of the different mechanisms to fluid redistribution cannot be evaluated. A key issue, affecting the volume of fluid redistributed by dilatancy pumping, is whether cyclic dilatational strains are localised to the material within fault zones or whether they extend over broad regions of the surrounding crust.

More than one of the dilatancy pumping mechanisms may contribute to fluid redistribution in any particular circumstance. To assess their relative contribution in different tectonic settings, it is necessary to consider the coupled variation of both shear and mean stress during fault loading (Fig. 2). In the case of a thrust fault, any dilatancy related to increasing shear stress during fault loading is opposed by the coupled rise in mean stress, whilst any postfailure tendency for crack closure from reduced shear stress is counteracted by lowered mean stress. In the case of a normal fault, however, development of dilatancy during loading is favoured by both the increasing shear stress and the coupled reduction in mean stress. Postfailure, reduced shear stress and increased mean stress both contribute to crack closure. On these arguments, dilatancy effects throughout the fault loading cycle should be more pronounced in extensional stress regimes.

$\Delta P_f$-dependent *hydrofracture dilatancy* is widely developed in the near-surface within extensional and strike-slip tectonic settings, but also develops in the lower reaches of the seismogenic zone in compressional regimes that are overpressured to lithostatic levels as an accompaniment to extreme fault-valve action. Localisation of hydrofracture arrays to the immediate vicinity of fault zones suggests that the fault zones themselves are the principal conduits for migration of overpressured fluids (Cox *et al.* 1991).

**Rupture nucleation**

The simple Coulomb criterion for frictional reactivation (eqn. 1) has been shown to be applicable to several cases of induced seismicity occurring in the top few kilometers of the crust (Nicholson &

Wesson, 1990). Its validity may well extend throughout the seismogenic zone, but fault failure at depth may also be affected by other fluid-related processes such as subcritical crack growth by stress corrosion, which allows time-dependent failure at constant stress (Das & Scholz, 1981). However, when repeated rupturing accompanied by valving action occurs in fluid-overpressured crust, the period prefailure is a time of increasing fluid pressure as well as shear stress, both of which contribute to fault instability and rupture nucleation (Figs. 1 & 3). In the case of severely misoriented faults, frictional mechanics suggests that it is the accumulation of fluid pressure to meet the necessary condition, $P_f > \sigma_3$, that is the actual trigger for rupture nucleation (Sibson, 1990).

TABLE 1 - Postulated Dilatancy Mechanisms

| MECHANISM | PRIMARY SENSITIVITY | REFERENCE |
|---|---|---|
| HIGH-STRESS MICROCRACK DILATANCY - at high $(\sigma_1 - \sigma_3)$ | $\Delta\tau$ | Nur, 1972<br>Scholz *et al.*, 1973 |
| LOW-STRESS MICROCRACK DILATANCY - from subcritical crack growth at low $(\sigma_1 - \sigma_3)$ | $\Delta\tau$ | Crampin *et al.*, 1984 |
| 'SAND-PILE' DILATANCY - under low $(\sigma_1 - \sigma_3)$ | $\Delta\tau$ | Nur, 1975 |
| EXISTING JOINT/FRACTURE DILATANCY | $\Delta\tau$ &/or $\Delta\bar{\sigma}$ ? | Nur, 1975 |
| $\Delta\bar{\sigma}$ COMPACTIVE EFFECTS - in highly fractured or porous material | $\Delta\bar{\sigma}$ | Sibson, 1991 |
| HYDROFRACTURE DILATANCY - under low $(\sigma_1 - \sigma_3)$ & high $P_f$ | $\Delta P_f$ | Sibson, 1981 |
| GRAIN-SCALE PARTICULATE FLOW - under high $P_f$ and low $\bar{\sigma}'$ | $\Delta P_f$ | Cox & Etheridge, 1989 |

**Rupture propagation**

There is geological evidence for several fluid-related effects linked to rupture propagation in fluid-saturated crust. Rapid slip transfer across dilational fault jogs and bends leads to abrupt localized reductions in fluid pressure and hydraulic implosion of wallrock into cavity space (Sibson, 1985).

Despite high anticipated levels of power dissipation during seismic slip ($\geq$10 MW/m$^2$), pseudotachylite friction-melt is comparatively rare in exhumed fault zones. The feedback interaction between temperature and pore-fluid pressure (Sibson, 1973; Lachenruch & Sass, 1980; Mase & Smith, 1987) may account for the scarcity of this material. Provided the fault zone contains comparatively impermeable material, the first increments of frictionally generated heat during slip are sufficient to boost fluid pressure and lower kinetic shear resistance to an extent that further heat generation is minimized. 'Clastic' dikes of gouge leading off slip surfaces may be

diagnostic of these transient fluid overpressures. Development of transient fluid overpressures behind the rupture front, before loss through hydrofracturing, provides one mechanism for the drastic slip-weakening and healing processes inferred to accompany rupturing by Heaton (1990).

**Earthquake swarms**

Hill (1977) attributed earthquake swarms to the migration of fluids through a 'honeycomb mesh' of interlinked shear and extensional fractures. In such systems, the passage of hydrothermal fluids triggers a cascade of minor shear ruptures accounting for the high b-value character of swarm activity. At other than very shallow depths, swarm activity is probably diagnostic of fluid overpressures. Field exposures document the existence of such fault/fracture meshes, best developed in extensional/transtensional tectonic settings, over a broad range of scales.

**Rupture arrest and aftershock activity**

Rupturing of an irregular fault surface leads to abrupt postfailure changes in mean stress localised around the structural irregularities (Segall & Pollard, 1980), with redistribution of fluids from areas of raised to areas of lowered mean stress (Nur & Booker, 1972; Li *et al.* 1987). Fluids are driven out of compressional jogs and bends where mean stress increases postfailure, while the most intense fluid influx, coupled with aftershock activity, is concentrated in regions of sharply reduced mean stress such as dilational jogs and bends. At these dilational sites, rapid slip transfer during rupture propagation causes abrupt local reductions in fluid pressure below ambient (hydrostatic?) values. In some instances, the induced suctional forces may lead to rupture arrest (Sibson 1985). Differing rock mass permeabilities can potentially account for variations in the time-dependence of aftershock sequences.

Dilational fault jogs and bends thus act essentially as *suction pumps* and are often characterised by multiply recemented wallrock breccias resulting from repeated hydraulic implosion, providing evidence for abrupt imbalances in fluid pressure. Larger dilational structures typically comprise a fault/fracture mesh of extension veins, implosion breccias, and subsidiary shears.

**Controls on earthquake recurrence**

The cycling of fluid pressures through the earthquake cycle as a consequence of fault-valve action may greatly influence recurrence behaviour (Fig. 3). Rather than being determined solely by the accumulation of shear stress to a "fault strength" that stays constant with time, as in the strictly periodic and time-predictable recurrence models, both shear stress and fault strength change through the interseismic period so that recurrence is determined by the intersection of rising shear stress with decreasing fault strength as fluid pressures reaccumulate.

**General Fault Zone Model**

A range of models accounting for the generation and maintenance of fluid overpressures in transcrustal fault zones have recently been proposed to explain the apparent weakness of the San Andreas and other major fault zones. A key issue in the different models is whether the high fluid pressures are derived from the fault zone acting as a migratory conduit for overpressured fluids (Sibson, 1990; Rice, 1992), or whether the overpressures are continually regenerated from essentially the same fluid volume during cyclical loading (Byerlee, 1990, 1993; Sleep & Blanpied, 1992). The development of extensive hydrothermal veining in fault zones, especially the gold-quartz mineralisation precipitated at structural levels corresponding to the lower half of the

seismogenic zone, provides geological evidence that fault zones, in at least some circumstances, are acting as migratory conduits for the passage of substantial fluid volumes (Cox *et al.* 1991).

If fluid overpressures and associated fault-valve action leading to fluid pressure cycling are as widespread as evidence is beginning to suggest, then fluid pressure gradients within transcrustal fault zones must be regarded as time-dependent, affecting the shear resistance profiles derived from rheological modelling. Figure 4 is an attempt to illustrate the effects on rheology and fault strength for a transcrustal fault zone that are likely to arise from fluid pressure cycling associated with fault-valve action. Both the depth and amplitude of the peak shear resistance become time-dependent. In addition, the integrated strength of the fault zone is at a minimum prefailure and reaches its maximum value at the end of the postfailure discharge phase, before self-sealing occurs and fluid pressure starts to reaccumulate.

## Conclusions

Aqueous fluids play a role in rupture nucleation, propagation, and arrest, and in the time-dependent character of aftershock and swarm activity. Much seismic faulting occurs in fluid-overpressured crust. Faulting then leads to complex interactions between stress cycling, the creation and destruction of permeability, and fluid flow, that need to be incorporated into general models of fault zone rheology and strength. Much more detailed information is needed on values of permeability in and around fault zones, the distribution of impermeable barriers, the localization of fluid overpressures, and the relative effectiveness of the various dilatancy pumping mechanisms. A principal goal must be increased understanding of the chemical and physical factors controlling the creation and destruction of permeability at different crustal levels.

## References


Batzle, M.L. & Simmons, G. 1977. Geothermal systems, rocks, fluids, fractures. *Am. Geophys. Un. Mon.* 20, 233-242.

Boullier, A-M. & Robert, F. 1992. Paleoseismic events recorded in Archean gold-quartz vein networks, Val d'Or, Abitibi, Quebec. *J. Struct. Geol.* 14, 161-179.

Byerlee, J.D. 1990. Friction, overpressure and fault-normal compression. *Geophys. Res. Lett.* 17, 2109-2112.

Byerlee, J.D. 1993. Model for episodic flow of high-pressure water in fault zones before earthquakes. *Geology* 21, 303-306.

Chester, F.M., Evans, J.P. & Biegel, R.L. 1993. Internal structure and weakening mechanisms of the San Andreas Fault. *J. Geophys. Res.* 98, 771-786.

Coombs, D.S. 1993. Dehydration veins in diagenetic and very-low-grade metamorphic rocks: features of the crustal seismogenic zone and their significance to mineral facies. *J. Metamorph. Geol.* 11, 389-399.

Cox, S.F. & Etheridge, M.A. 1989. Coupled grain-scale dilatancy and mass transfer during deformation at high fluid pressures: examples from Mount Lyell, Tasmania. *J. Struct. Geol.* 11, 147-162

Cox S.F., Wall, V.J., Etheridge, M.A. & Potter, T.F. 1991. Deformational and metamorphic processes in the formation of mesothermal vein-hosted gold deposits - examples from the Lachlan Fold Belt in central Victoria, Australia. *Ore Geol. Rev.* 6, 391-423.

Crampin, S., Evans, R. & Atkinson, B.K. 1984. Earthquake prediction: a new physical basis. *Geophys. J. R. astr. Soc.* 76, 147-156.

Das, S. & Scholz, C.H. 1981. Theory of time-dependent rupture in the earth. *J. Geophys. Res.* 86, 6038-6051.

Grant, N.T., Banks, D.A., McCaig, A.M. & Yardley, B.W. 1990. Chemistry, source, and behavior of fluids involved in Alpine thrusting of the Central Pyrenees. *J. Geophys. Res.* 95, 9123-9131.

Grindley, G.W. & Browne, P.R.L. 1976 Structural and hydrological factors controlling the permeability of some hot-water geothermal fields. *In*, Proc. 2nd United Nations Symposium on Development and Use of Geothermal Resources, New York, 377-386.

Heaton, T.H. 1990. Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Phys. Earth Planet. Int.* 64, 1-20.

Hill, D.P. 1977. A model for earthquake swarms. *J. Geophys. Res.* 82, 1347-1352.

Hubbert, M.K.& Rubey, W.W. 1959. Role of fluid pressure in the mechanics of overthrust faulting. *Bull. Geol. Soc. Am.* 70, 115-205.

Lachenbruch, A.H. 1980. Frictional heating, fluid pressure, and the resistance to fault motion. *J. Geophys. Res.* 85, 6097-6112.

Li, V.C., Seale, S.H. & Cao, T. 1987. Postseismic stress and pore pressure readjustment and aftershock distributions. *Tectonophysics* 144, 37-54.

Mase, C.W. & Smith, L. 1987. Effects of frictional heating on the thermal, hydrological, and mechanical response of a fault. *J. Geophys. Res.* 92, 6249-6272.

Morrow, C., Lockner, D., Moore, D. & Byerlee, J. 1981. Permeability of granite in a temperature gradient. *J. Geophys. Res.* 86, 3002-3008.

Mullis, J. 1988. Rapid subsidence and upthrusting in the Northern Appenines, deduced by fluid inclusion studies in quartz crystals from Porretta Terme. *Schweiz. Mineral. Petrogr. Mitt.* 68, 157-170.

Nicholson, C. & Wesson, R.L. 1990. Earthquake hazard associated with deep well injection - a report to the U.S. Environmental Protection Agency. *U.S. Geol. Surv. Bull.* 1951.

Nur, A. 1972. Dilatancy, pore fluids, and premonitory variations of $t_s/t_p$ travel times. *Bull. Seism. Soc. Am.* 62, 1217-1222.

Nur, A. 1975. A note on the constitutive law for dilatancy. *Pure Appl. Geophys.* 133, 197-206.

Nur, A. & Booker J.R., 1972. Aftershocks caused by pore-fluid flow? *Science* 175, 885-887.

Parry, W.T. & Bruhn, R.L. 1990. Fluid pressure transients on seismogenic normal faults. *Tectonophysics* 179, 335-344.

Rice, J.R. 1992. Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault. *In,* Evans, B. & Wong, T-F. (eds.) *Fault Mechanics and Transport Properties of Rocks*, Academic Press, New York, 475-503.

Scholz, C.H., Sykes, L.R., Aggarwal, Y.P. 1973. Earthquake prediction: a physical basis. *Science* 181, 803-810.

Segall, P. & Pollard, D.D. 1980. Mechanics of discontinuous faults. *J. Geophys. Res.* 85, 4337-4350.

Sibson, R.H. 1973. Interactions between temperature and pore-fluid pressure during earthquake faulting - a mechanism for partial or total stress relief. *Nature Phys. Sci.* 243, 66-68.

Sibson, R.H. 1981. Controls on low-stress hydrofracture dilatancy in thrust, wrench, and normal fault terrains. *Nature* 289, 665-667.

Sibson, R.H. 1985. Stopping of earthquake ruptures at dilational fault jogs. *Nature* 316, 248-251.

Sibson, R.H. 1990. Rupture nucleation on unfavorably oriented faults. *Bull. Seism. Soc. Am.* 80, 1580-1604.

Sibson, R.H. 1991. Loading of faults to failure. *Bull. Seism. Soc. Am.* 81, 2493-2497.

Sibson, R.H. 1992. Implications of fault-valve behaviour for rupture nucleation and recurrence. *Tectonophysics* 211, 283-293.

Sleep, N.H. & Blanpied, M.L. 1992. Creep, compaction, and the weak rheology of major faults. *Nature* 359, 687-692.

Smith, D.L. & Evans, B. 1984. Diffusional crack healing in quartz. *J. Geophys. Res.* 89, 4125-4135.

Sprunt, E.S. & Nur, A. 1977. Destruction of porosity by pressure solution. *Geophysics* 42, 726-741.

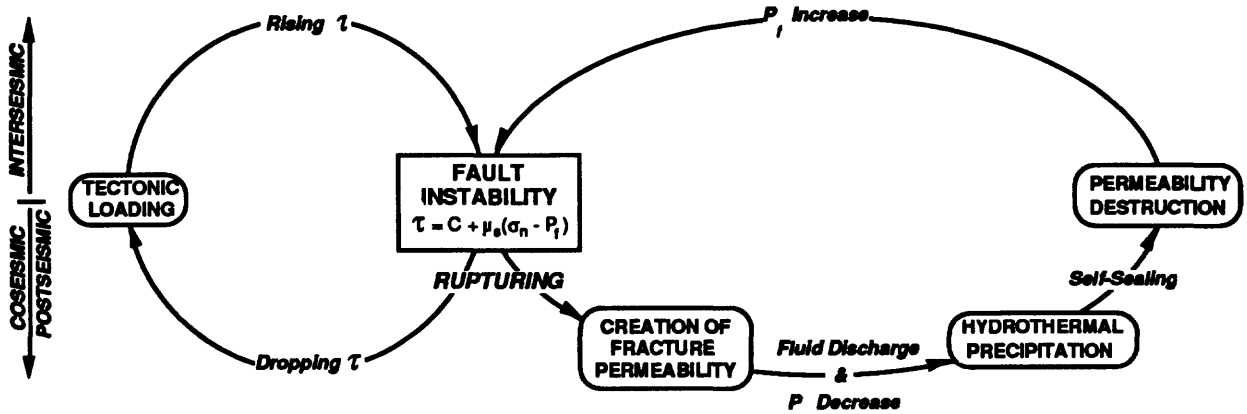



Figure 1 - Synoptic diagram of *fault-valve* activity, illustrating the dependence of fault failure on both stress and fluid pressure cycling (after Sibson, 1992).

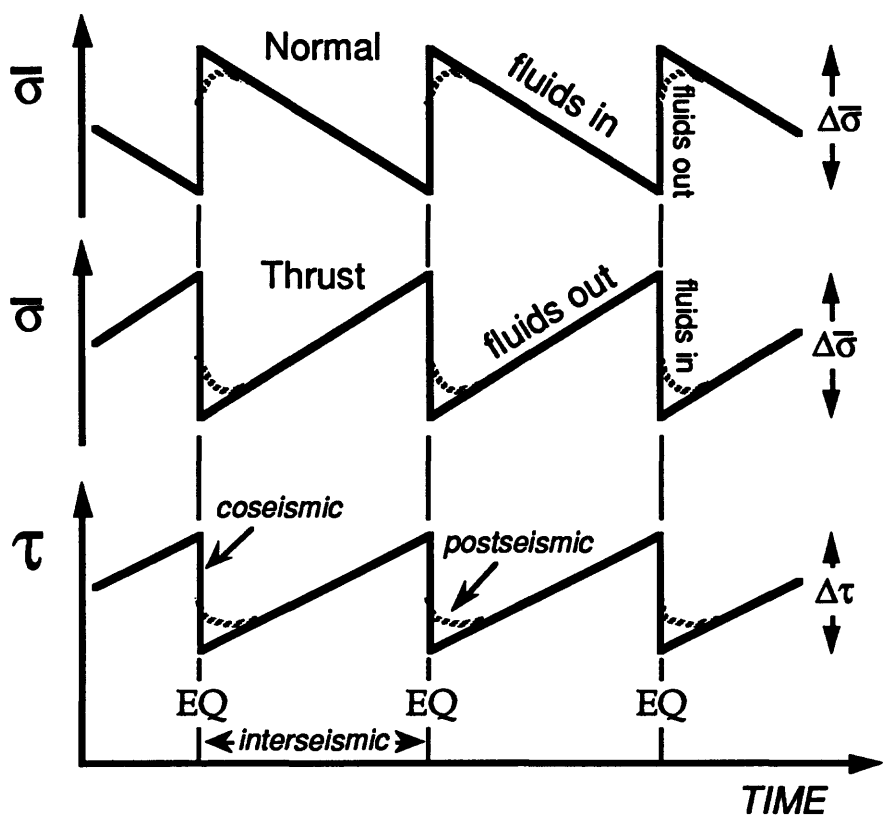



Figure 2 - Variations in mean stress ($\bar{\sigma}$) during the cyclic accumulation and release of shear stress ($\tau$) on thrust and normal faults that are optimally oriented for frictional reactivation with the vertical stress assumed to stay constant. Fluid movements in and out of the response zone around the fault are indicated in relation to different phases of the earthquake stress cycle.

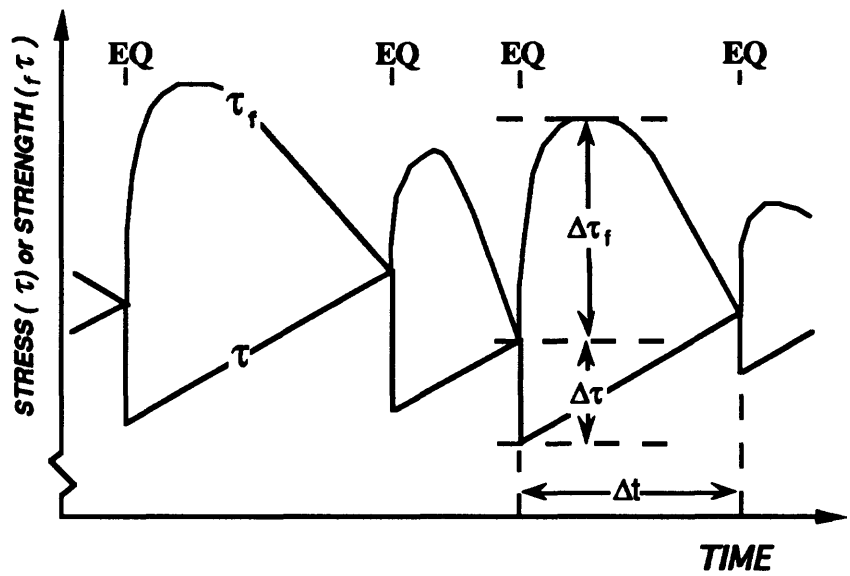



Figure 3 - General recurrence model involving fault-valve activity (after Sibson, 1992). Postfailure discharge leads to time-dependent fault strengthening until self-sealing occurs. Reaccumulation of fluid pressure then causes progressive fault weakening. Rate of shear stress accumulation stays constant, but the recurrence interval, $\Delta t$, is controlled by the time-dependence of both $\tau_f$ and $\tau$.

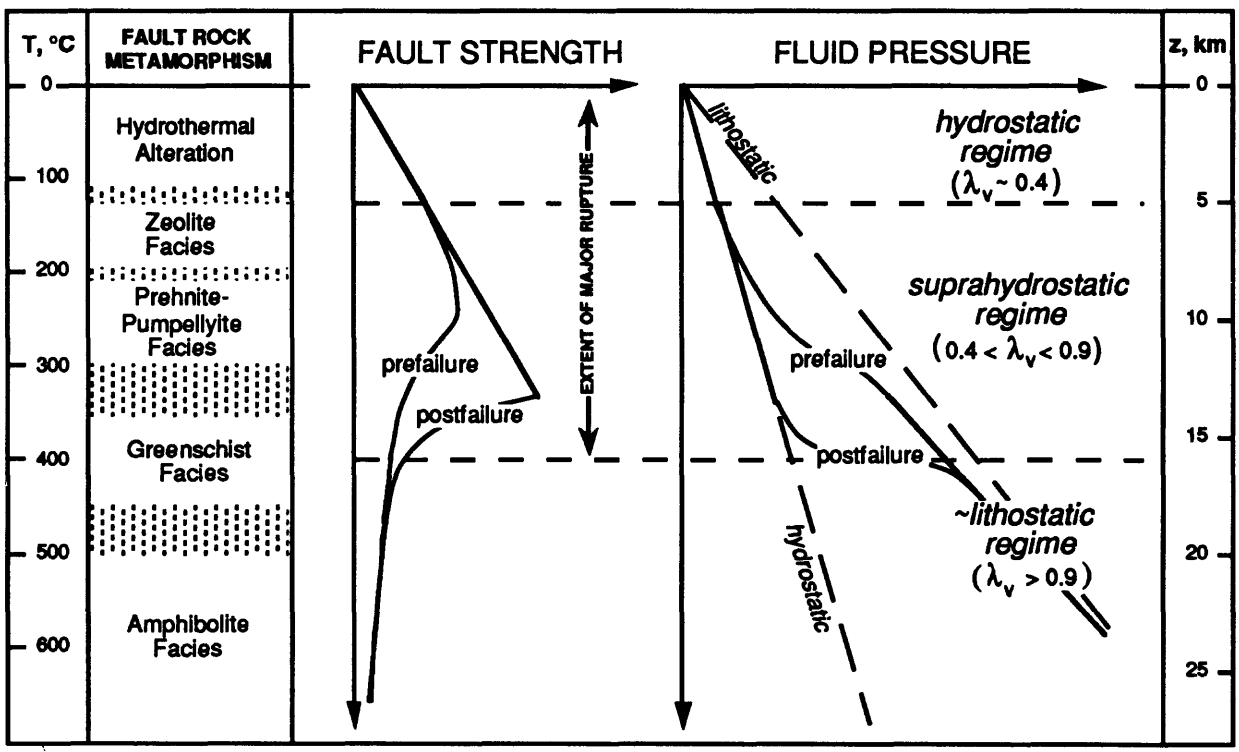



Figure 4 - General schematic model of a fault zone incorporating fault-valve behavior.

# Fluids and Faulting: Suggestive Evidence and Crucial Tests at Parkfield

P.A. Johnson and T.V. McEvilly

*Seismographic Station and*
*Center for Computational Seismology, Lawrence Berkeley Laboratory,*
*University of California, Berkeley*

## Abstract

At Parkfield, along the locked-to-creeping transition of the San Andreas fault in central California, source mechanisms and the development of some earthquake sequences reveal interesting features with important implications for the mechanics of fault failure. Temporal and spatial patterns of some clustered earthquakes reveal a structure in which the events start at a point and proceed to spread outwards, usually horizontally in one or both directions along the fault zone. Some clusters occur in the form of *en echelon* trends at an angle to the San Andreas fault, suggestive of the tension gash arrays observed by structural geologists and perhaps of wing cracks. Seismic P- and S-wave radiation and near-field moment tensor inversions of clustered earthquakes permit a measurable component of tensile (CLVD) motion, while there is some evidence that non-clustered earthquakes are pure shear failure.

These observations of the spatial-temporal evolution of cluster patch rupture and the source mechanisms of both clustered and non-clustered earthquakes can be construed as support for the growing number of conceptual models which involve crustal fluids at pore-fluid pressures that fluctuate between low and high throughout the earthquake cycle. These models predict certain patterns for the development of rupture and types of earthquake source mechanisms that can be used to test competing models for rupture initiation.

We examine a plausible model for fluid-controlled seismicity at Parkfield in which microearthquake clusters are proposed to represent small localized 'pods' of fluctuating pore-fluid pressure. Throughout the earthquake cycle, the pressure within a pod gradually increases, due perhaps to the inexorable volumetric strain in the fault zone, to decreasing permeability in fluid flow paths, heating effects, dehydration, etc., until the pressure approaches lithostatic. At this point, the cluster patch is weaker than the surrounding fault zone, and it fails in a process initiated with a central fluid-driven event (not unlike a hydrofracture) that triggers shear failure on the pre-stressed fault surface and is followed by a diffusion of failure outwards in conventional shear-source microearthquakes, to distances typically as far as 200 m but in some cases as much as two kilometers from the initiating failure. Non-clustered earthquakes, in this failure model, occur at 'conventional' strong asperities by mechanisms that do not require a major role for fluids. We present some observations of microearthquakes at Parkfield that support this conceptual model.

## 1. Introduction

When the U.S. Army was injecting waste fluid at high pressure into a deep well at the Rocky Mountain Arsenal near Denver, Colorado during the 1960s, a rash of local earthquakes triggered serious consideration of the possible role of fluids in the faulting process. A field experiment (Raleigh *et al.*, 1976) to test the causal role of the decrease in effective normal stress on the faults at depth, due to the increased pore-fluid pressure, in the triggering of the earthquakes showed conclusively that the occurrence of earthquakes can be controlled by modulating the fluid pressure in a fault zone.

At about the same time, measurements of heat flow across the San Andreas fault were failing to reveal a significant anomaly centered over the fault (Henyey, 1968; Henyey and Wasserburg, 1971; Lachenbruch and Sass, 1973). This was contrary to what was expected from laboratory rock mechanics experiments of friction and failure. In a fractured rock, as we may envision the crust surrounding the San Andreas fault to be, frictional sliding will occur on the fractures at a level of stress below the fracture strength of intact rock. Laboratory experiments for precut rock samples indicated that this frictional strength is about 100 MPa, under pressure and temperature conditions believed to be obtained in the upper crust (Byerlee and Brace, 1968, 1969; Byerlee, 1970). However, the absence of any detectable heat flow anomaly over the fault constrains the fault shear stress to be an order of magnitude lower, in the range of 10 to 25 MPa (Brune *et al.*, 1969), implying an unexpectedly low coefficient of friction on the fault (less than 0.1), much lower than the 0.6 to 0.8 of Byerlee's law.

During the Cajon Pass Drilling Project in the 1980's, the maximum principal compressive stress at depth near the San Andreas fault was found to be nearly perpendicular to the fault, implying a very small resolved shear stress on the fault (Zoback *et al.*, 1987). This finding is in conflict with laboratory observations of the mechanics of intact rock, in which the maximum principal stress is oriented at a small angle (25 to 30 degrees) to the slip plane; however, a long-lived large-displacement structure with a kinematically-controlled stress field, such as the San Andreas fault, can slip under a maximum principal stress oriented at a steeper angle to the plane. The characteristics of various geological structures found along and adjacent to the San Andreas fault offer strong support for a steeper, near-normal orientation of the maximum principal stress near the fault. Numerous anticlines, synclines, and thrust faults strike parallel to the fault, indicating a significant amount of compression perpendicular to the fault. Other stress indicators, such as borehole breakouts, provide evidence that the maximum principal stress direction is oriented at a high angle to the fault.

The absence of a fault-centered heat flow anomaly and the fault-normal orientation of the maximum principal compressive stress along the San Andreas fault suggest that the fault is weak. Two possible explanations are that either the fault zone is composed of very weak (low friction) material, presumably fault gouge, or fluids under high pressure are present within the fault zone. The weakest gouge mineral likely present along the San Andreas fault in significant quantity is montmorillonite clay, but this has a coefficient of friction of $\geq 0.2$ (Radney and

Byerlee, 1988), double the value constrained by the heat flow data. It is highly improbable that this clay is present in sufficient quantity to dominate the frictional strength of the fault and result in an average fault friction coefficient of around 0.2.

Anomalous fluid conditions remain as the most plausible mechanism for the weak fault. Fluids within the fault zone at sufficient volume and pressure result in low effective normal stress on the fault, allowing the fault to slip under low resolved shear stress and explaining the lack of a heat flow anomaly.

In this paper, we first review some of the evidence for fluids at high pore pressure in the crust and some of the models in which fluids under high pressure within fault zones exert significant control on fault slip and the nucleation of earthquakes. Building on this background, we use new high-resolution microearthquake observations at Parkfield to present some suggestive evidence for the involvement of fluids in the faulting process there, along the transitional stretch of the San Andreas fault from locked to creeping behavior that produces repeating M6 earthquakes.

## 2. Evidence and Observations of Fluids and High Pore Pressures in the Crust

There is evidence for the presence of fractures in the upper 10 or 20 km of Earth's crust. To depths of 5 to 10 km, there are both direct observations of fractures in mines and drill holes, and inferred evidence, based on electrical and fluid conductivity measurements, which argues for interconnecting, fluid-filled fractures to these depths (see Brace, 1972, for a summary). A crustal permeability study (Brace, 1980) argues for high conductivities down to 8 km, and deep resistivity measurements (Nekut *et al.*, 1977) push the case for water-filled, interconnected crustal fractures down to 20 km.

Evidence for fluid pressures well above hydrostatic at crustal depths greater than a few kilometers is plentiful and varied. For example, there are borehole measurements of high fluid pressures in various tectonic regimes, including sedimentary basins, accretionary prisms, and crystalline basement (Fertl *et al.*, 1976; Yerkes *et al.*, 1990; Davis *et al.*, 1983; and Kerr, 1984). Fluid pressure increases from hydrostatic to 90 percent of lithostatic between 2 and 5 km depth in sedimentary basins adjacent to major active strands of the San Andreas fault system (Berry, 1973; Yeats, 1983; Yerkes *et al.*, 1985; Sibson, 1990a). Berry (1973) describes an extensive (250-315 km in length) geographic region in which pore-fluid pressures within the Franciscan and Great Valley sediments reach near-lithostatic values. It is generally expected that in regions of prograde metamorphism at depth, fluid pressures approach lithostatic (Etheridge *et al.*, 1983).

## 3. Hypotheses and Models for Fluid-Controlled Fault Slip

Several hypotheses and associated phenomenological models have been put forth in which fluids and high pore pressures within the crust play a critical role in the mechanics of fault slip and the nucleation of earthquakes. Reported

observations in the early 1970s of temporal variations in seismic wave velocities prior to earthquakes featured a decrease in the ratio of P-wave to S-wave velocities (Vp/Vs), followed by an increase preceding a number of earthquakes (Nersesov *et al.*, 1969; Sadovsky *et al.*, 1972; Whitcomb *et al.*, 1973; Scholz *et al.*, 1973). To explain the phenomenon by known laboratory experimental results, Nur (1972) proposed a dilatancy-diffusion model, in which the increasing stress in the nucleation zone resulted in dilatancy of the rock, causing a decrease in Vp, and thus in Vp/Vs. As water flows into the dilated zone, Vp/Vs increases as pore pressure rises until it reduces the effective normal stress to the point that rupture takes place.

Another mechanism for earthquake failure involving fluids was put forth by Raleigh (1977). Following suggestions that frictional sliding leads to melting on the fault surface (Jaeger, 1962; McKenzie and Brune, 1972; Richards, 1976), Raleigh calculated the frictional heating during fault motions and concluded that the resulting rise in temperature is sufficient for the dehydration of clays and other hydrous phases in fault gouge. The water released during dehydration increases the pore pressure which can result in earthquake rupture. Lachenbruch (1980) considered pore fluid expansion due to frictional heating on the fault.

Sibson (e.g., 1992) has proposed that faults behave as fluid-pressure activated valves, allowing hydrofracturing of the rock in the fault zone. In his model, faults during the interseismic period are highly impermeable and act as pressure seals until the pressure exceeds the strength of the rock and failure can occur. The seismic rupture that takes place dramatically increases the permeability of the fault zone, so that the fault becomes a channel for fluid flow and the discharge of fluids from depth.

Models proposed by Byerlee (1990, 1993) and Rice (1992) call for distinct differences between the state of stress within a weak fault zone and that within the adjacent country rock and explain how this spatially varying stress tensor is compatible with the requisite high pore pressures that explain the absolute and relative weakness of the San Andreas fault. In their shear-failure models the magnitude of the minimum principal stress is everywhere greater than the high fluid pressure within the fault zone, precluding hydrofracture.

We investigate the above suite of proposed fluid-controlling processes in faulting using high-resolution microearthquake observations. An extensive data set of this type has been collected along the San Andreas fault zone in central California, and we now explore the models with those data.

## 4. Hypothesis Testing Using the High-Resolution Data at Parkfield

At Parkfield, in central California, the 10-station borehole-installed network of 3-component sensors has produced 7 years of data that may help to discriminate among the various proposed theories for the role of fluids in controlling fault-zone properties and failure conditions. A 30-km segment of the fault in the vicinity of the presumed M6 nucleation zone is being studied in fine detail. In this section, we review the following three lines of investigation which are pertinent to the role of

fluids in the faulting process:

1. three-dimensional P- and S-wave velocity structures;
2. temporal and spatial patterns of seismicity; and,
3. earthquake source mechanisms.

For each of the above, we describe the general observations and discuss or propose fluid mechanical models for seismicity and earthquake nucleation, in the context of the observations. For the latter two topics, we present our ongoing investigations for testing the proposed models of fluid involvement in the mechanics of fault slip at Parkfield.

## 4.1 Velocity Structure

### *4.1.1 Observations*

The three-dimensional velocity structure at Parkfield has been determined by the joint inversion for hypocenter locations and independent P- and S-wave models using the high-resolution borehole network data at Parkfield (Michelini and McEvilly, 1991) and by a similar inversion for the P-wave velocity structure using Calnet data (Eberhart-Phillips and Michael, 1993). The combined P and S velocity structure reveals two salient features (Figure 1; the coordinate system of this and all subsequent maps is defined with respect to the velocity inversion model: $x$ is position (in km) northeast across the fault, $y$ is the position (in km) northwest along the fault, and $z$ is model depth (elevation) (in km), which is one kilometer shallower than true depth. The point $(x,y,z) = (0,0,-1)$ is the location of the epicenter of the 1966 mainshock.) First, in the locked segment of the fault (to the southeast), a relatively high P- and S-velocity body, with normal Vp/Vs ratio, is present southwest of the San Andreas fault below 5 km depth. The P and S velocities of 6.6 and 3.6 km/s appear too high to be the granitic Salinian block. This high-velocity body seems to control the mode of deformation, the seismicity pattern and the extent of rupture in larger events, a role similar to that proposed for a deep high-velocity body at Loma Prieta (Michelini, 1991; Foxall, 1992; Foxall *et al.*, 1993). The P-wave velocity model of Eberhart-Phillips and Michael (1993) also includes a high velocity body in the vicinity of that found by Michelini and McEvilly (1991).

The second notable feature in the 3-D model has relevance for the role of fluids in crustal processes. This is the region of high Vp/Vs ratio (about 1.9), localized within the fault zone at a depth of 8 km and extending from the 1966 mainshock focus southeastward about 5 km into the 1966 rupture zone. The dimensions of the high Vp/Vs anomaly are close to the model resolution from the inversion process. Within these limits, the zone is clearly a flattened ellipsoid within the fault plane, elongated in the southeast-northwest direction of the fault. Maximum dimensions (based on the 1.9 ratio contour) of the zone appear to be 5-7 km by 3 km within the fault zone and perhaps 1 km wide normal to the fault. The anomaly lies within the region of depressed P- and S-wave velocities characteristic of the

fault zone but seems to be caused by a relatively greater reduction of the S-wave velocity in the localized zone. Eberhart-Phillips and Michael (1993) also find a region of low P velocity which is quite extensive (30 km x 50 km) and has its greatest intensity localized in the region of the high Vp/Vs anomaly of Michelini and McEvilly (1991).

The high Vp/Vs anomaly is located in the same general area that has shown a progressive velocity change in the coda of waveforms in the Vibroseis monitoring experiment (Karageorgi *et al.*, 1992) and has become a major focus of intense study at Parkfield. It is tempting to associate it with anomalous conditions related to the nucleation process of the M6 earthquake.

### *4.1.2 Models: Possible Explanations for the High Vp/Vs Anomaly*

*Computational artifact.* It is possible that the high Vp/Vs anomaly is simply an artifact of the computation, but formal resolution in the numerical inversion problem is reasonably good in that region compared to elsewhere in the model volume (Michelini and McEvilly, 1991). Thus, there is no reason to suspect a localized systematic error.

*Shear-wave anisotropy.* S-wave anisotropy is known to exist in the fault zone at shallower depths (Varian well VSP survey and Vibroseis monitoring program, e.g., Daley and McEvilly, 1990; Karageorgi *et al.*, 1992) and is best explained as a fault-zone fabric effect, with a horizontal axis of symmetry normal to the fault. To produce the localized S-wave velocity reduction by selective birefringence requires a contrived combination of S-wave polarizations and propagation directions unique to the anomalous zone. In addition, this peculiarity must apply to S waves from sources below the anomalous zone along the entire southeast-northwest section of the fault zone, comprising a wide range of ray paths. Furthermore, less than 10% anisotropy is seen at the shallower depths, whereas the anomalous S-wave velocity suppression is more than 15%. Of course, there may well be a component of S-wave anisotropy contributing to the anomaly, but it is not likely to explain both the Vp and the Vs anomalies.

*Extensive saturated fractures.* This is a real possibility, if open fluid-filled fractures can exist at depths of 5 to 10 km within fault-zone materials (a deep equivalent of gouge, yet having lower velocity than normal crustal constituents at those depths). The anomalous Vp/Vs zone at Parkfield is the presumed nucleation region of an M6 earthquake probably late in its recurrence cycle and, as such, may be highly fractured due to dilatancy.

*High pore-fluid pressure (with or without fractures).* Our favored explanation for the anomalously high Vp/Vs zone is that the region is overpressured. Eberhart-Phillips and Michael (1993) also suggest high fluid pressure as the cause of their low P-wave velocity region, within which the high Vp/Vs zone resides. They point out that very low resistivity, lack of a low gravity signature, and highly magnetic material (probably serpentinite) above the low-velocity feature are also consistent with high fluid pressure in this area. Porosity of the rock may exist as fractures or as diffuse intergranular space (or both). Elevated pore pressure will significantly

decrease the S-wave velocity, but will cause only a small decrease in the P-wave velocity, resulting in a high Vp/Vs (Nur, 1972). High pore pressure has been documented along the San Andreas fault at shallow depths (less than about 4-5 km, e.g., Berry, 1973; Sibson, 1990a). At seismogenic depths, plausible mechanisms for increasing the pore pressure include permeability and porosity reduction as a result of the inexorable volumetric strain of relatively weak fault-zone material due to fault-normal compression, decreased permeability in fluid flow paths, thermal expansion of fluids due to frictional heating, dehydration of clay minerals, etc. The possibility of such mechanisms operating at observable rates in the Parkfield nucleation zone has made the Vp/Vs anomaly a prime target in the monitoring program for premonitory change, using both clustered microearthquakes and the Vibroseis as highly repeatable sources of P- and S-wave illumination of the anomalous zone.

## 4.2 Spatial and Temporal Characteristics of Seismicity

### *4.2.1 Observations*

Seismicity at Parkfield from 1987 through 1992 is distributed along the central 30 km stretch of the fault containing the Middle Mountain nucleation zone of the M6 earthquakes (Figure 2). The distribution of earthquakes during this period is consistent with the long-term pattern of seismicity evident since 1970, as described by Bakun and Lindh (1985) and Poley *et al.* (1987). Hypocenters are limited mostly between depths of 2 and 12 km. The general patterns of seismicity, including relatively quiet 'gaps', areas of high concentration of earthquakes, and regions of more diffuse activity, has been largely stable from year to year (Figure 3) (Nadeau *et al.*, 1993). However, there appears to be some evidence for migration of activity northwestward from $y = -3$ km into the nucleation zone of the M6 and M4.7, beginning in 1989 and culminating in the October 1992 M4.7 and the November 1993 M4.8 sequences that filled the gap from $y = -3$ to $y = +3$ km through the M6 nucleation zone at 8-10 km depth. The patterns of both the background seismicity and the dense concentrations of activity, as observed by the high-resolution network since 1987, seem related to various features of the velocity model (for example, the high velocity body and the high Vp/Vs region), to the 1966 hypocenter and rupture zone, and to the overall fault zone transition from locked to creeping behavior across the nucleation zone (Foxall, 1992).

Very fine-scale concentrations of earthquakes also occur, almost anywhere along the seismically active part of the fault. These clusters typically have dimensions of 100 to 200 m and consist of anywhere from 2 to 20 or more earthquakes, with interevent times from 2 seconds to 4 years or more (Nadeau *et al.*, 1993). About 80 clusters have been identified in the period 1987-89, and although they occupy only a fraction of the total fault surface (about 1-2%) that will likely slip in the M6 event, they represent about 50% of the earthquakes recorded and a significant amount of the seismic energy. For example, the M4.7 of October 1992 is the mainshock member of a cluster with 12 members identified to date, having magnitudes as small as -0.5 (Johnson and McEvilly, 1992). This fine-scale

clustering in microearthquake occurrence is a fundamental characteristic of the fault-zone dynamics at Parkfield and bears on the mode of seismic strain release along that stretch of the fault.

Using a cross-correlation/cross-spectrum technique, relative timing of the P and S arrivals among cluster members can be obtained to sub-sample timing (i.e., to less than 2 ms), resulting in relative locations accurate to about 5 to 10 meters. This has provided a very detailed picture of the rupture process in these fault zone patches, within which successive seismic failure occurs with an interevent spacing of meters to tens of meters, spanning an ultimate extent of 100 to 200 meters for most clusters of highly similar events.

After applying this high-resolution relocation process to some of the clusters, we find intriguing geometrical structure at sub-cluster size, suggestive of intricate slip processes. In several of the clusters analyzed to date, we find what appear to be *en echelon* lineations of hypocenters. Figure 4 shows two such clusters. The first (Figure 4a) appears to consist of two *en echelon* strands 50 m apart and 100 m long, as well as a third roughly equidimensional (within the resolution of the locations) sub-cluster. The cluster shown in Figure 4b may show two *en echelon* segments about 70 m apart and 100 to 200 m long, although the sparse distribution within this particular cluster precludes more definitive statements. In Figure 4, the *en echelon* segments are rotated about 40 degrees clockwise from the N45W local strike of the San Andreas fault, which lies parallel to the $y$–axis.

### *4.2.2 Models of Fluid Involvement in Clustered Microearthquakes*

In one hypothetical model for fluid-controlled fault slip at Parkfield, the clusters of concentrated seismic slip represent localized 'pods' of high fluid pressure. We can imagine the fault zone to consist of a complex distribution of separate hydrologic domains, much as Byerlee (1993) proposed. Observed in deep oil reservoirs, fluid domains are seen to be disconnected, seal-bounded fluid compartments of various dimensions, porosities, and pore-fluid pressures. An analogous model for seismicity at Parkfield, both clustered and non-clustered, calls for the discretization of the fault zone into fluid domains. The small clusters of earthquakes, especially those consisting of nearly coincident, highly similar members within small volumes (less than about 200 m), have potential for discerning fine-scale patterns of fault zone slip and yielding evidence for the existence of domains and how they change through space and time. The repetitive nature of the sources concentrated within the clusters (on time scales ranging from minutes to days to years) suggests a modulating local effective stress. In this model fluid injection, generation, or compression causes increasing pore pressure within a cluster pod until the fluid is forced out of the pod as a hydrofracture. The initiating hydrofracture may be limited to the very first break of the fault patch, and the event may grow immediately into a conventional shear failure. Alternatively, an initiating event in a sequence may be substantially tensile failure. In either case, the resulting fault slip creates increased fracture permeability and porosity, draining fluids from the patch, increasing the effective normal stress and inhibiting further slip. During the interseismic

period, fracture permeability is reduced by, for example, mineral precipitation sealing the compartment or by volumetric compressive strain, or perhaps the fluid pressure is increased by thermal expansion or dehydration, allowing failure to occur in a cyclic manner until the anomalous fluid concentration dissipates or migrates to another, hydrologically more attractive, patch on the fault.

A second model reverses the above pathology of cluster patch failure, defining clusters as strong, relatively dry patches distributed throughout an otherwise fluid-weakened fault zone that is largely aseismic at the microearthquake scale. Stated briefly these two models are:

Model 1. Earthquake clusters occur in weak (saturated and overpressured) patches and the surrounding fault zone is strong (relatively dry);

Model 2. Earthquake clusters occur in strong (relatively dry) patches and the surrounding fault zone is weak (saturated and overpressured).

Schematic diagrams of these two models are presented in Figure 5. The first model predicts that as the fault zone undergoes increasing strain, pore pressure in the cluster patches will increase until the effective normal stress is small enough that the patch can slip. This would force the fluid from the patch into the relatively more dry and stronger surrounding regions of the fault zone, a process which would manifest itself as an outward propagation with time of the earthquakes in a cluster patch, with some tensile (hydrofracture) component, at least in an initial event of a sequence.

There are two variations of Model 2. If in the second model the mechanism for patch rupture is fluid-dominated (as depicted in Figure 5), then as the fault zone accumulates strain, fluid in the surrounding fault zone would be forced into the relatively drier cluster patches, decreasing the effective normal stress on the patch progressively inwards toward the center of the patch, allowing the patch to slip under shear as it does so. However, this model does not require a dominant fluid mechanism. In a second variation on Model 2, the patches represent conventional strong asperities which fail without the influence of effective normal stress reduction through increased fluid pressure. The patch weakens by some other mechanism (e.g., stress corrosion, strain weakening, etc.) and subsequently fails dynamically as a common shear failure. This variation of Model 2 is also the basic model for the failure of non-clustered earthquakes at Parkfield.

These contrasting models for fault zone slip can be tested with the fine-scale microearthquake occurrence patterns within individual clusters of events (next section). In addition, there should be detectable differences between the source mechanisms of the initiating cluster events of Model 1 and those of both variations of Model 2. Model 1 failure would involve some component of hydrofracture, while patch failure in either variation of Model 2 would be accomplished by a common double-couple shear mechanism. Non-clustered earthquakes would also fail with double-couple shear motion (see Section 4.3). There may be a difference in

stress drop between the three models of failure. Depending on the fraction of shear stress relieved (which would depend on the slip-locking process) and on the area involved in a given slip amplitude (which may depend on the size of the pod or asperity), the model mechanisms may differ in stress drop.

### *4.2.3 Testing the Models: Observations of the Evolution of Cluster Patch Rupture*

In an attempt to test the models, we have been investigating geometrical, temporal and mechanistic properties of the clustered earthquakes and their surrounding environment within the fault zone. For the present work, we consider any spatially and temporally related set of earthquakes as clusters. With such a definition, the clusters we have admitted are anywhere from 100 m to a couple kilometers in spatial extent, but of short duration (hours to days).

We start by considering the spatial and temporal evolution of cluster failure. The locations we use are those from the Parkfield high-resolution catalog (located using the 3-D velocity model). Repicking and careful relocations (fixed-station set, same phase weights, etc.) result in a shift of the entire pattern by an amount less than about a kilometer, but the relative locations retain their patterns to 10-20 m. It may be possible to improve incrementally the full cluster data set by such work, which is underway.

Figure 6 shows the activity starting with the M4.7 mainshock in October 1992, followed by its immediate and some of its later aftershocks, plotted as a function of time $t$ versus each of $x$, $y$, and $z$, the three orthogonal directions defining the fault plane (as in Figures 1 and 2). The cluster in Figure 6 is the largest-scale cluster we have examined to date, and it shows the general features we have observed in numerous clusters. Clusters seem to start at some point on the fault and with subsequent earthquakes they tend to spread out from this point with time, usually horizontally along the fault, but sometimes also in depth or across-fault. Many clusters then experience a burst of activity at a later time back in the center of the patch, near where the initial rupture took place. In the cluster shown in Figure 6, the first earthquakes occurred at time $t \approx 92.804$ (decimal year; October 20, 1992 at 05:28 UTC) in a tight group at a position of about $y = -3$ km along the fault, $x = -2$ km across the fault, and a depth of 10 km ($z = -9$ km); the initiating event was the M4.7 mainshock. From here, activity spread outward in both directions along the fault (+ and $-$ $y$ directions, i.e., northwest and southeast) and shallower and deeper. This activity was followed by a burst of aftershocks at $t \approx 92.821$ (October 26) near the first-ruptured region; this set of aftershocks repeated the general pattern by propagating simultaneously to the northwest and southeast ($\pm y$) along the fault. Figure 7 shows this second burst in more detail. The sequence terminated about a day later at $t \approx 92.823$ (October 27) with a few earthquakes in the center of the patch that was ruptured the previous day.

Figures 8, 9, 10, and 11 present four other clusters which show the same types of rupture behavior with time. In some cases, the first-ruptured portion of the patch remains active (Figure 8); other clusters propagate only outwards but culminate with a burst in the vicinity of initial rupture (Figures 7 and 9), and the two clusters

shown in Figures 10 and 11 grow monotonically outwards. The earthquakes of these sequences usually spread horizontally along the fault in one or both directions, but sometimes they propagate in depth also. The cluster shown in Figure 10 grows downwards and to the southwest, perpendicular to the main fault, for about 1 km, perhaps along a conjugate fracture plane (an *en echelon* trend?), at a high angle to the San Andreas fault.

The propagation characteristics of these clusters do not reliably demonstrate the flow of fluid outwards from the patch centers as overpressurization and subsequent fault slip takes place, but it is an admissable process. An alternative mechanism, of course, is that the clustered earthquakes are triggered by the outward propagation of stress pulses due to the release and redistribution of strain energy from the previous earthquakes.

## 4.3 Source Mechanisms

### *4.3.1 Motivation and Models*

Our exploration of these hypotheses for the role of fluids in the nucleation of both the M6 and the small clusters of earthquakes is motivated in large part by the likely effect of fluids in shaping the velocity structure, in particular the region of high Vp/Vs. In addition, the spatial and temporal evolution of clustered events and their repetitive nature are suggestive of the diffusion of water and microearthquakes outwards from the cluster center, and of a cycling in effective stress within the cluster patch, respectively. Furthermore, while most of the first motion mechanisms determined for earthquakes along the San Andreas fault tend to be strike-slip shear, we have found that the initiating events of two of the clusters (Figures 7 and 11) indicate mostly normal faulting motion, which may be indicative of a tensional (hydrofracture) component of motion, although one of those events could also be fit by a strike-slip mechanism. These characteristics argue for the first (weak patch, high pore-fluid pressure) model of earthquake cluster failure (Model 1; Figure 5). This model is likely to involve a failure process which is initiated with a central fluid-driven high stress drop hydrofracture event. The spatial patterns in the clusters suggest a complexity in the faulting process at Parkfield that may involve a range of slip type. A careful investigation of source mechanisms and their distribution is clearly neccessary.

We are also motivated in this exploratory research by numerous other lines of evidence for the possible existence of non-shear earthquake mechanisms. Laboratory experiments have demonstrated that shear failure in intact rock is accompanied by tensile failure in the form of wing cracks, which develop near the tips of the shear crack; initially and nearest the shear crack tip, the wing crack is oriented at a high angle to the shear crack but farther from the main crack, the wing crack rotates and becomes oriented parallel to the direction of maximum principal stress (e.g., Nemat-Nasser and Horii, 1982). Fracture theory predicts crack tip stresses consistent with the numerous laboratory and field observations of wing cracks (Pollard and Segall, 1987). Development of a macroscopic shear plane occurs by the

coalescence of tensile cracks (Hallbauer *et al.*, 1973; Nemat-Nasser and Horii, 1982). Studies of joint formation in granite show clearly that voids open at seismogenic depths (e.g., Martel *et al.*, 1988; Martel and Pollard, 1989). Figure 12 shows the wing crack model and the *en echelon* arrays of tension gashes oft observed by structural geologists, and their analogy to the *en echelon* lineations along the San Andreas fault at Parkfield as imaged by earthquake epicenters.

Seismological evidence for non-shear fault mechanisms includes the observation that a substantial fraction of earthquakes induced at depth in massive granite appear to have a significant tensional component (Feignier and Young, 1991; Gibowicz *et al.*, 1991), and some microearthquakes at Long Valley caldera have been shown to involve a substantial amount of non-double-couple body wave radiation (Peppin and Johnson, 1994).

These observations and theoretical results, coupled with the growing evidence for the major role of high pore pressure in facilitating slip within fault zones (e.g., Sibson, 1992; Rice, 1992; Nur and Walder, 1992; Byerlee, 1993), prompts a search for tensional-failure events indicative of local hydrofracturing.

### *4.3.2 Testing the Models*

In analyzing the data for the subtle evidence of a non-double-couple source component using small earthquakes (magnitudes less than 1.5 or 2), conventional first-motion focal mechanisms will not suffice because of the limited close-in azimuthal coverage and lack of high-gain high-resolution instruments. However, with high quality data from modern high-sensitivity networks concentrated in small regions, it is now possible to try novel methods of determining source properties of earthquakes over a large magnitude range. We now explore this topic using the Parkfield data.

Following Peppin and Johnson (1994) and Johnson *et al.* (1992), we compute S-to-P spectral ratios in the frequency band below 5 Hz for three-component data at a number of recording sites at a range of azimuths. We choose reasonable models for shear and tensile sources and compare the predicted values of the ratios to the observations.

The model adopted for the shear source is taken from the double-couple solution for the strike, dip, and rake of the San Andreas fault and of its auxiliary plane. These fault parameters were combined with the angles-of-incidence and azimuth information obtained by location through the three-dimensional velocity model. Because of the symmetry of double-couple sources, each choice of strike, dip, and rake gives the same predicted radiation patterns of S/P. We calculate the S/P far-field spectral ratio for the shear source from equations 4.33 of Aki and Richards (1980).

For the tensional source radiation pattern, we chose the model that follows from fracture mechanical theory, which is in agreement with the observations of numerous laboratory experiments of failure. Thus, the orientation of the wing cracks trends toward the direction of maximum compressive stress (Nemat-Nasser

and Horii, 1982). In this model, two solutions are obtained for the S/P spectral ratio because the wing cracks can be formed by either of the two possible shear fault planes (i.e., the San Andreas and its auxiliary plane) and associated slip vectors. For the tension source, the S/P spectral ratio is computed using equations of 3.12 and 3.13 of Walter (1991).

We illustrate the analysis in Figure 13, which shows the results we have been obtaining for the mechanisms of clustered and non-clustered earthquakes. It is possible to interpret the events as shear, tensional, or mixed corresponding to the degree to which the observations fit the predictions. In this manner, the spectral ratios of the cluster event in Figure 13 and those of other cluster events suggest that these earthquakes involve a substantial non-double-couple component, while the spectral ratios for the non-clustered event agree well with a double-couple mechanism.

Another promising method we have been investigating for the determination of source mechanisms is near-field moment tensor inversion (Uhrhammer, 1992). From a single three-component broadband station in the near-field (within about 35 km, depending on the source size and signal-to-noise ratio), it is possible to recover a robust estimate of the full, unconstrained seismic moment tensor, including all six independent components of the seismic source. Using a least-squares procedure in the time domain, the inversion finds the complete unconstrained moment tensor that best fits the observations, using a halfspace model derived from the traveltimes and hypocenter. The best-fitting double-couple solution is also calculated. Of the near-field terms visible in the displacement seismogram (the ramp and offset), the ramp provides the most constraint on the moment tensor. It is the low frequency nature of the ramp that allows the halfspace model to work so well; a layered halfspace model can provide a better fit to details of the waveforms, but it does not significantly improve the resolvability of the moment tensor components (Uhrhammer, 1992).

We have been using U.C. Berkeley's Broadband Digital Seismic Network (BDSN) station PKD1 at Parkfield (80 samples per second) for the inversion of Parkfield earthquakes and have been obtaining some interesting results. Figure 14 shows examples of near-field moment tensors obtained for two Parkfield events (the $M_W$ 3.7 on February 15, 1993 and the $M_W$ 4.9 on April 4, 1993). Synthetic seismograms are shown along with the observed data. The near-field ramps are clearly visible on all three components (vertical, radial, and transverse) for both earthquakes. The orthogonal lines of the mechanisms represent the best-fitting double-couple solution; shaded regions represent the full unconstrained solution. Both events have a significant non-shear (CLVD) component (17% and 15%, respectively). The degree to which the unconstrained moment tensor (shaded region) overlaps or fails to meet at the intersection of the double-couple nodal planes is controlled by the interaction of the CLVD and isotropic components of the decomposition of the tensor. These two non-double-couple components can work together or against each other, so without looking at the individual components, it is impossible to say whether the mechanism is tensile or compressive by observing the full solution. We have broken the tensor of the February event into its three components and the result indicates a component of tensile motion for this event. Compared to

other regions along the San Andreas fault, a higher proportion of the near-field moment tensors of events at Parkfield are showing a substantial amount of CLVD component (Uhrhammer, pers. comm., 1993).

The results of these source mechanism studies show that events in some clusters appear to have a measurable tensile failure in their mechanisms, suggesting the possibility that we may finally be observing at a scale in which the fault rupture initiation may be studied over lengths of meters and times of milliseconds, and where modern instrumentation allows unambiguous identification of non-shear fault slip.

## 5. Discussion

Parkfield is proving to be a region of complex faulting at all scales. The *en echelon* patterns, the spatial and temporal evolution of cluster patch rupture, the diffusion-like propagation of events outwards from the patch and its repetitive nature, and the evidence for non-shear tensile failure all lend support to Model 1, in which earthquake cluster patches are high pressure pods waiting for a break.

Model 1 suggests that the initiating events of the clusters are hydrofractures. This may be observable under the right conditions (as our observations indicate is possible), but it is likely to be difficult to see in most cases. If fluids are involved and all the cluster events have identical waveforms, then either all the events are hydrofractures, or the hydrofrac component of the initiating event is overwhelmed by its later shear slip history, which dominates the waveform and accounts for most of the energy release. Another possibility is that the initiating events may not always be identified with a cluster. For example, if members of a cluster are identified by the similarity of their waveforms but the waveform of the initiating event is strikingly different due to a substantial tensile component, it would not be identified as the initiating event of the cluster, with our waveform-based criterion.

Diffusion velocities of cluster sequence event propagation are highest near the pod center (about 170 km/day) and decrease as the process moves outward until velocities of about 1-10 km/day are reached.

Fluids at high pressures almost certainly exist within the fault zone and may play a critical role in the microseismicity and the nucleation of the impending M6 event. The nucleation and failure mechanism of the M6 Parkfield earthquake may be similar to that proposed for the clusters. The 5-7 km by 3 km tabular region of high Vp/Vs ratio and low seismicity (until the October 1992 and November 1993 sequences) near the 1966 hypocenter (Foxall, 1992) defines a zone of distinct physical properties in the presumed M6 nucleation zone. Our hypothesis is that it is due to localized high pore-fluid pressure in either the 'normal' porosity extant throughout most of the fault zone, or within dilatancy-related fractures in the nucleation zone. The failure process culminating in the M6 event must first effect slip on the high Vp/Vs region (the 'usual' Parkfield M5 foreshock?) and when this 'megapod' grows, perhaps in a scaled up version of the cluster slip process, it quickly breaks the entire vertical section (2 to 12 km depth) of the fault zone, with the resulting 'runaway' slip being the M6 event.

## 6. Summary and Conclusions

Modern high-resolution borehole-emplaced seismographic networks concentrated in regions of intense seismicity are revealing a new level of detail in the pathology of major fault zones (e.g., Anza and Parkfield, CA) and other active complex structures (e.g., Long Valley caldera, The Geysers geothermal reservoir). Yielding relative hypocenter locations accurate to a few tens of meters, 1-2 km resolution in independently estimated P- and S-wave velocity structures (and Vp/Vs ratio), and the fine-scale spatial and temporal distribution of seismicity, these observational tools may finally provide long-sought direct evidence for fluid-controlled fault failure.

In this paper we have shown some details from the Parkfield earthquake studies that have implications for the possible role of fluids in the dynamics of the fault-zone. The combined approach of both the fine-scale view of coupled geometrical/temporal evolution of cluster failure along with the investigation of earthquake source mechanisms, especially for small earthquakes (magnitudes less than 1.5 or 2), may expose finally the suspected and likely important role of fluids in fault-zone failure. The anomalous Vp/Vs within the deep fault zone seems to indicate the presence of high fluid pressures. The earthquake clusters may delineate localized pods of high fluid pressure. Spatial-temporal migration within and among clusters reflects a stress diffusion process. Source mechanisms appear to contain a non-shear component indicative of hydrofracturing, consistent with a model involving induced seismic breakout of high pressure pods along the fault zone.

## Acknowledgements

The Parkfield High-Resolution Seismic Network functions through close cooperation among researchers at the University of California at Berkeley, Lawrence Berkeley Laboratory, Duke University, and the U.S. Geological Survey. USGS provides financial support for this research through NEHRP award 14-08-001-G1703. Data processing was done at the Center for Computational Seismology at LBL, which is operated by the University of California for the U.S. Department of Energy under contract No. DE-AC03-76SF00098. We thank Lind Gee for a thoughtful review and Bob Uhrhammer for re-running his near-field moment tensor solutions and his help in preparing Figure 14.

## References

Aki, K. and Richards, P.G., 1980. *Quantitative Seismology: Theory and Methods*, W.H. Freeman, New York.

Bakun, W. and Lindh, A., 1985. The Parkfield, California prediction experiment, *Earthq. Predict. Res.* **3**, 285-304.

Berry, F.A.F., 1973. High fluid potentials in the California Coast Ranges and their tectonic significance, *Bull. Am. Assoc. Petrol. Geol.* **57**, 1219-49.

Brace, W.F., 1972. Pore pressure in geophysics, in *Flow and Fracture of Rocks, Geophys. Monogr. Ser. 16*, edited by H.C. Heard, I.Y. Borg, N.L. Carter, and C.B. Raleigh, 265-273, AGU, Washington, D.C.

Brace, W.F., 1980. Permeability of crystalline and argillaceous rocks: Status and problems, *Int. J. Rock Mech. Mining Sci.* **17**, 876-893.

Brune, J.N., Henyey, T.L., and Roy, R.F., 1969. Heat flow, stress, and rate of slip along the San Andreas fault, California, *J. Geophys. Res.* 74, 3821-3827.

Byerlee, J., 1970. Static and kinetic friction of granite at high normal stress: *Int. J. Rock Mech. Mining Sci.* **7**, 6, 577-582.

Byerlee, J., 1990. Friction, overpressure and fault normal compression, *Geophys. Res. Lett.* **17**, 2109-2112.

Byerlee, J., 1993. A model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology* **21**, 4, 303-306.

Byerlee, J.D., and Brace, W.F., 1968. Stick-slip, stable sliding, and earthquakes--effect of rock type, pressure, strain rate, and stiffness, *J. Geophys. Res.* **73**, 18, 6031-6037.

Byerlee, J.D., and Brace, W.F., 1969. High-pressure mechanical instability in rocks, *Science* **164**, 3880, 713-715.

Daley, T.M. and T.V. McEvilly, 1990. Shear wave anisotropy in the Parkfield Varian Well VSP, *Bull. Seism. Soc. Am.* **80**, 857-869.

Davis, D., Suppe, J., and Dahlen, F.A., 1983. Mechanics of fold-and-thrust belts and accretionary wedges, *J. Geophys. Res.* **88**, 1153-1172.

Eberhart-Phillips, D. and Michael, A.J., 1993. Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California, *J. Geophys. Res.* **98**, 15,737-15,758.

Etheridge, M.A., Wall, V.J., and Vernon, R.H., 1983. The role of the fluid phase

during regional metamorphism and deformation, *J. metamorph. Geol.* **1**, 205-226.

Feignier, B. and Young, R.P., 1991. Moment tensor inversion of induced microseismic events: Source mechanism investigation in the $-4 < M < -1$ moment magnitude range, *EOS, Trans. Am. Geophys. Un.* **72**, 292.

Fertl, W.H., Chilingarian, G.V. and Rieke, H.H., 1976. *Abnormal Formation Pressures*: Elsevier, Amsterdam, 382 pp.

Foxall, W., 1992. Heterogeneous slip and rupture models of the San Andreas fault zone based upon three-dimensional earthquake tomography, Ph.D. Dissertation, University of California, Berkeley, 168 pp.

Foxall, W., Michelini, A. and McEvilly, T.V., 1993. Earthquake tomography of the southern Santa Cruz mountains: Control of fault rupture by lithological heterogeneity of the San Andreas fault zone, *J. Geophys. Res.* (in press).

Gibowicz, S.J., Young, R.P., Talebi, S., and Rawlence, D.J., 1991. Source parameters of seismic events at the underground research laboratory in Manitoba, Canada: Scaling relations for events with moment magnitude smaller than −2, *Bull. Seism. Soc. Am.* **81**, 1157-1182.

Hallbauer, D.K., Wagner, H., and Cook, N.G.W., 1973. Some observations concerning the microscopic and mechanical behavior of quartzite specimens in stiff, triaxial compression tests, *Int. J. Rock Mech. Mining Sci. Geomech. Abstr.* **10**, 713.

Henyey, T.L., 1968. Heat flow near major strike-slip faults in central and southern California, Ph.D. Dissertation, California Institute of Technology.

Henyey, T.L., and Wasserburg, G.J., 1971. Heat flow near major strike-slip faults in California, *J. Geophys. Res.* **76**, 7924-7946.

Jaeger, J.C., 1962. *Elasticity, Fracture and Flow*: John Wiley & Sons, Inc., New York.

Johnson, P.A. and McEvilly, T.V., 1992. Magnitude scaling and source parameter variations of microearthquakes at Parkfield, *EOS Trans. Am. Geophys. Un.* **73**, 398.

Johnson, P.A., Antolik, M., Peppin, W.A., Foxall, W., and McEvilly, T.V. 1992. Structure and rupture processes of the fault zone at Parkfield from the spectra of microearthquakes, *Seism. Res. Lett.* **63**, 75.

Karageorgi, E., Clymer, R., and McEvilly, T.V., 1992. Seismological studies at Parkfield. II. Search for temporal variations in wave propagation using Vibroseis, *Bull. Seism. Soc. Am.* **82**, 1388-1415.

Kerr, R.A., 1984. Continental drilling headed deeper, *Science* **224**, 1420.

Lachenbruch, A.H., 1980. Frictional heating, fluid pressure and the resistance to fault motion, *J. Geoph. Res.* **85**, 6097-6112.

Lachenbruch, A.H., and Sass, J.H., 1973. Thermo-mechanical aspects of the San Andreas fault system, in *Proceedings of the Conference on Tectonic Problems of the San Andreas fault*, Stanford University Press, Stanford, CA.

Martel, S.J., Pollard, D.D., and Segall, P., 1988. Development of simple strike-slip fault zones, Mount Abbot quadrangle, Sierra Nevada, California, *Bull. Geol. Soc. Am.* **100**, 1451-1465.

Martel, S.J. and Pollard, D.D., 1989. Mechanics of slip and fracture along small faults and simple strike-slip fault zones in granitic rock, *J. Geophys. Res.* **94**,

9417-9428.

McKenzie, D.P. and Brune, J.N., 1972. Melting of fault planes during large earthquakes, *Geophys. J. R. Astron. Soc.* **29**, 65-78.

Michelini, A., 1991. Fault zone structure determined through the analysis of earthquake arrival times, Ph.D. Dissertation, Univ. Calif., Berkeley, 191 pp.

Michelini, A. and McEvilly, T.V., 1991. Seismological studies at Parkfield: I. Simultaneous inversion for velocity structure and hypocenters using B-splines parameterization, *Bull. Seism. Soc. Am.* **81**, 524-552.

Nadeau, R., Antolik, M., Johnson, P.A., Foxall, W., and McEvilly, T.V., 1993. Seismological Studies at Parkfield III: Microearthquake clusters in the study of fault-zone dynamics, *Bull. Seism. Soc. Am.* (in press).

Nekut, A., Connerney, J.E.P., and Kuckes, A.F., 1977. Deep crustal electrical conductivity: Evidence for water in the lower crust, *Geophys. Res. Lett.* **4**, 6, 239-242.

Nemat-Nasser, S. and Horii, H., 1982. Compression-induced nonplanar crack extension with application to splitting, exfoliation, and rockburst, *J. Geophys. Res.* **87**, 6805-6821.

Nersesov, I.L., Semenov, A.N., and Simbireva, I.G., 1969. Space-time distribution of the travel time ratios of transverse and longitudinal waves in the Garm area, in *The Physical Basis of Foreshocks*, Akad. Nauk SSSR Publication.

Nur, A., 1972. Dilatancy, pore fluids, and premonitory variations of ts/tp travel times, *Bull. Seism. Soc. Am.* **62**, 1217-1222.

Nur, A. and Walder, J., 1992. Hydraulic pulses in the earth's crust, in *Fault Mechanics and Transport Properties of Rocks*: Academic Press.

Peppin, W.A. and Johnson, P.A., 1994. Low frequency S to P spectral ratios: Evidence for nonshear components in microearthquakes near Long Valley caldera, California, *Bull. Seism. Soc. Am.* (in press).

Poley, C.M., Lindh, A.G., Bakun, W.H., and Schulz, S.S., 1987. Temporal changes in microseismicity and creep near Parkfield, California, *Nature* **327**, 134-137.

Pollard, D.D. and Segall, P., 1987. Theoretical displacements and stresses near fractures in rock: With applications to faults, joints, veins, dikes, and solution surfaces, in *Fracture Mechanics of Rock*, edited by B.K. Atkinson, 277-349 pp., Academic Press, San Diego, Calif.

Radney, B. and Byerlee, J., 1988. Laboratory studies of the shear strength of montmorillonite, *EOS Trans. Am. Geophys. Un.* **69**, 1463.

Raleigh, C.B., 1977. Frictional heating, dehydration and earthquake stress drops, in *Proceedings of Conference II Experimental Studies of Rock Friction with Application to Earthquake Prediction*, 291-304, U.S. Geol. Surv., Menlo Park, CA.

Raleigh, C.B., Healy, J.H., and Bredehoeft, J.D., 1976. An experiment in earthquake control at Rangely, Colorado, *Science* **191**, 1230-1237.

Ramsay, J.G., 1967. *Folding and fracturing of rocks*, McGraw-Hill, New York.

Rice, J. R., 1992. Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault, in *Fault Mechanics and Transport Properties of Rocks*: Academic Press.

Richards, P.G., 1976. Dynamic motions near an earthquake fault: A three-dimensional solution, *Bull. Seism. Soc. Am.* **66**, 1-32.

Sadovsky, M.A., Nersesov, I.L., Nigmatullaev, S.K., Latynina, L.A., Lukk, A.A., Semenov, A.N., Simbireva, I.G., and Ulomov, V.I., 1972. The processes preceding strong earthquakes in some regions of Middle Asia, *Tectonophysics* **14**, 295-307.

Scholz, C.H., Sykes, L.R., and Aggarwal, Y.P., 1973. The physical basis for earthquake prediction, *Science* **181**, 803-810.

Segall, P. and Pollard, D.D., 1983. Nucleation and growth of strike slip faults in granite, *J. Geophys. Res.* **88**, 555-568.

Sibson, R.H., 1990a. Conditions for fault-valve behaviour, in Deformation mechanisms, rheology and tectonics, edited by R.J. Knipe and E.H. Rutter, *Geol. Soc. London Sp. Pub.* **54**, 15-28.

Sibson, R.H., 1990b. Rupture nucleation on unfavorably oriented faults, *Bull. Seism. Soc. Am.* **80**, 6, 1580-1604.

Sibson, R.H., 1992. Implications of fault-valve behaviour for rupture nucleation and recurrence, *Tectonophysics* **211**, 283-293.

Uhrhammer, R., 1992. Broadband near-field moment-tensor inversion, *EOS Trans. Am. Geophys. Un.* **73**, 375.

Walter, W.W., 1991. High Frequency Seismic Source Spectra from Earthquakes and Explosions, Ph.D. Dissertation, University of Nevada, Reno, April 1991.

Whitcomb, J.H., Garmany, J.D., and Anderson, D.L., 1973. Earthquake prediction: Variation of seismic velocities before the San Fernando earthquake, *Science* **180**, 632-635.

Yeats, R.S., 1983. Large-scale Quaternary detachments in Ventura basin, southern California, *J. Geophys. Res.* **88**, 569-583.

Yerkes, R.F., Levine, P. and Wentworth, C., 1985. Abnormally high fluid pressures in the region of the Coalinga earthquakes - a preliminary report, *U.S. Geol. Surv. Open File Rep. 85-44*, 344-375.

Yerkes, R.F., Levine, P., and Wentworth, C.M., 1990. Abnormally high fluid pressures in the region of the Coalinga earthquake sequence and their significance, in The Coalinga Earthquake of May 2, 1983, edited by M.J. Rymer and W.L. Ellsworth, *U.S. Geol. Surv. Prof. Pap. 1487*, 235-257.

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

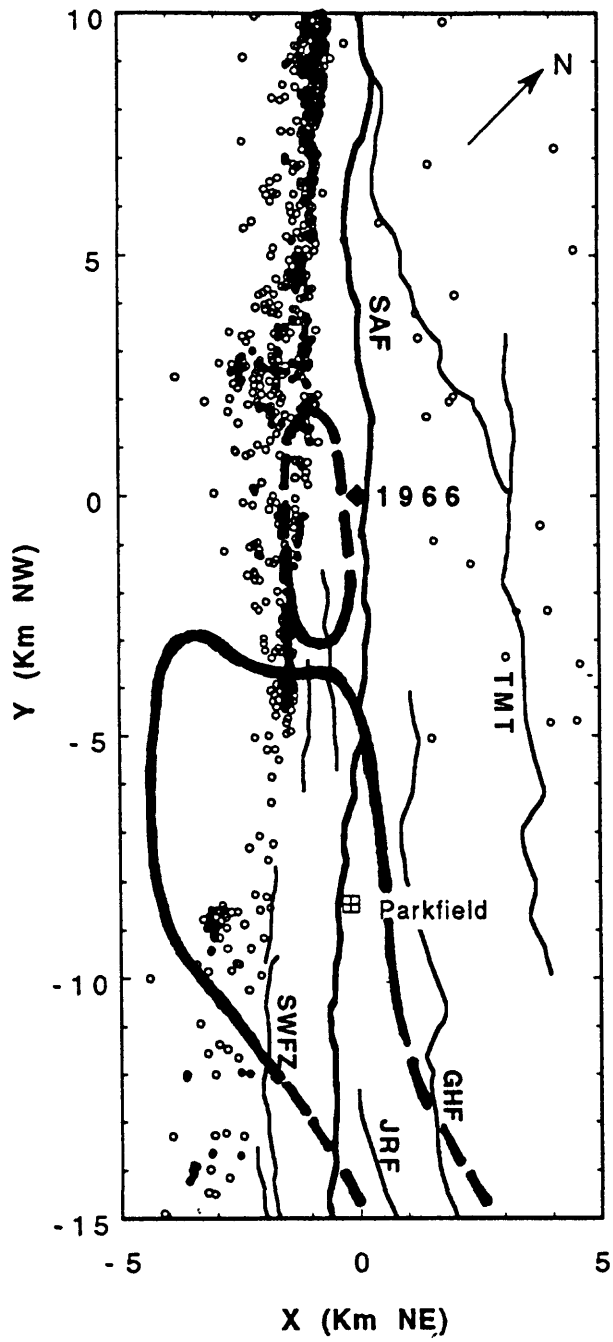



**Figure 1.** Map view of the deep (>5 km) high-velocity body ($-3 < y < -11$) and the deep (8 km depth) high Vp/Vs anomaly ($2 < y < -3$) (bold outlines). Also shown are the 1966 epicenter (solid diamond), clustered (solid dots) and other (open circles) background seismicity; and the town of Parkfield. Major faults: GHF, Gold Hill; JRF, Jack Ranch; SAF, San Andreas; SWFZ, Southwest Fracture Zone; TMT, Table Mountain thrust. Map is oriented with respect to the San Andreas fault, centered at the 1966 epicenter: $y$-axis is distance in kilometers northwest from the 1966 epicenter; $x$-axis is distance in kilometers northeast from the 1966 epicenter. *After Foxall, 1992.*

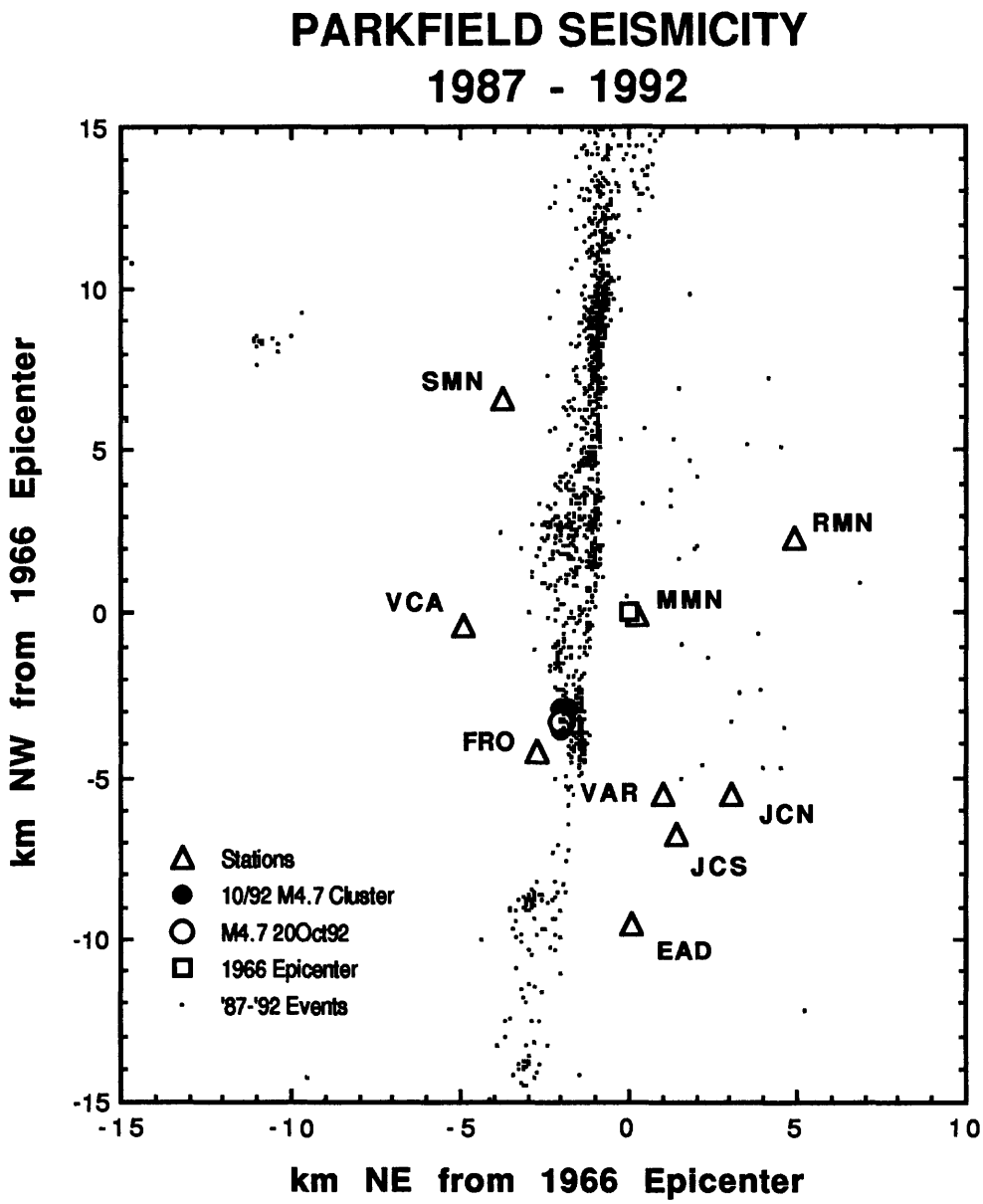



**Figure 2.** Map showing the Parkfield study area. Included are the seismographic stations of the borehole network (open triangles); microearthquake locations February 1987 - December 1992 (small dots); the 1966 M6 epicenter (open square over MMN, Middle Mountain); October 1992 M4.7 cluster (large solid dots); and the M4.7 October 20, 1992 mainshock (large open circle). Map is oriented with respect to the San Andreas fault, centered at the 1966 epicenter: $y$-axis is distance in kilometers northwest from the 1966 epicenter; $x$-axis is distance in kilometers northeast from the 1966 epicenter.

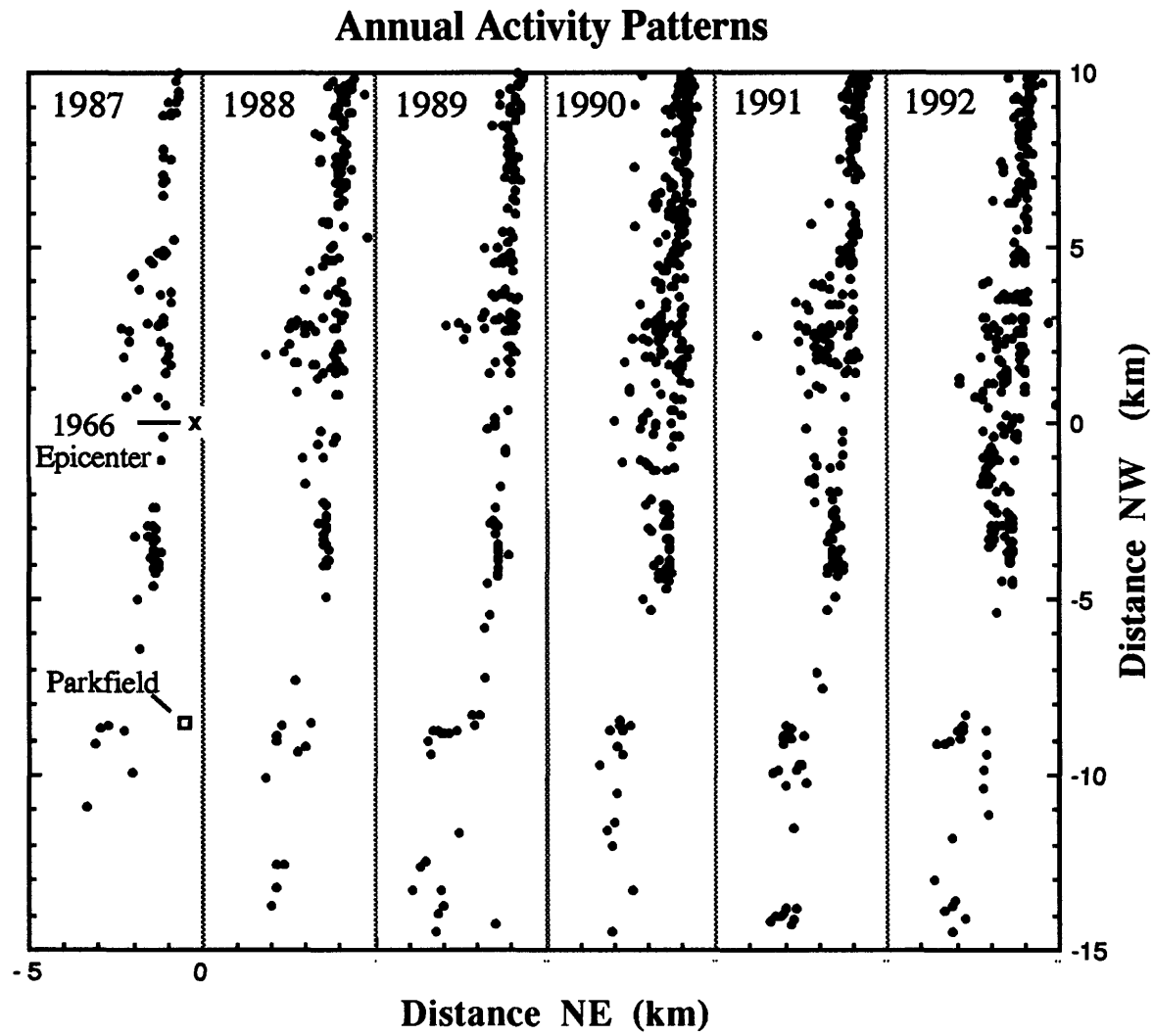



**Figure 3.** Strip maps of annual seismicity for the years 1987-1992. The 1966 epicenter (symbol x) and the town of Parkfield (open square) are shown. Maps are oriented with respect to the San Andreas fault, centered at the 1966 epicenter: $y$-axis is distance in kilometers northwest from the 1966 epicenter; $x$-axis is distance in kilometers northeast from the 1966 epicenter.

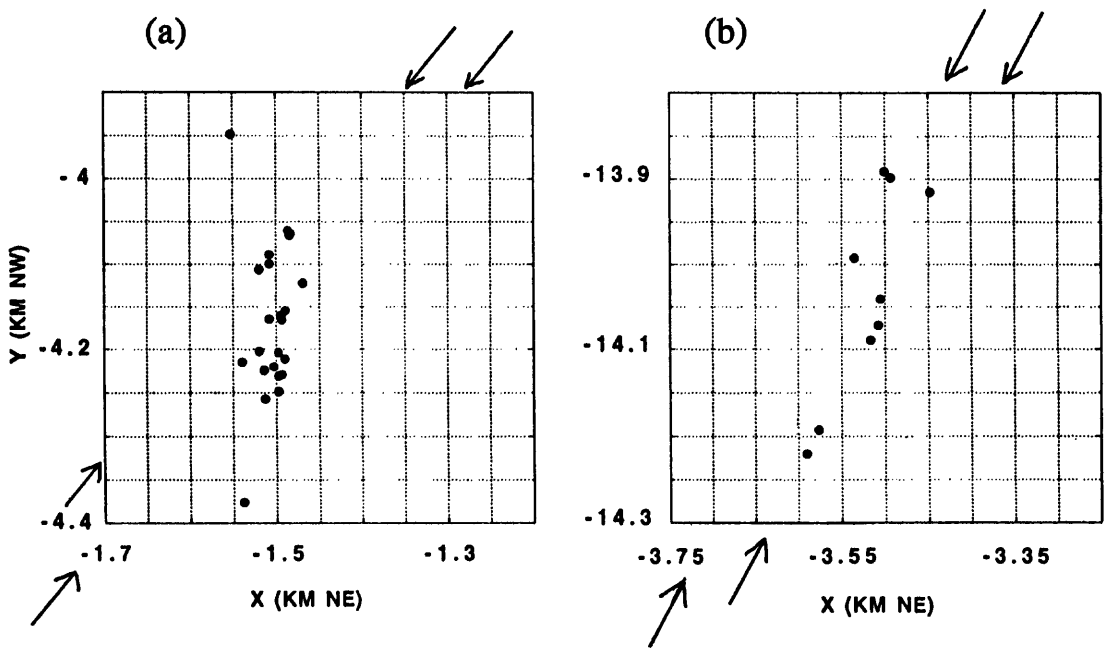
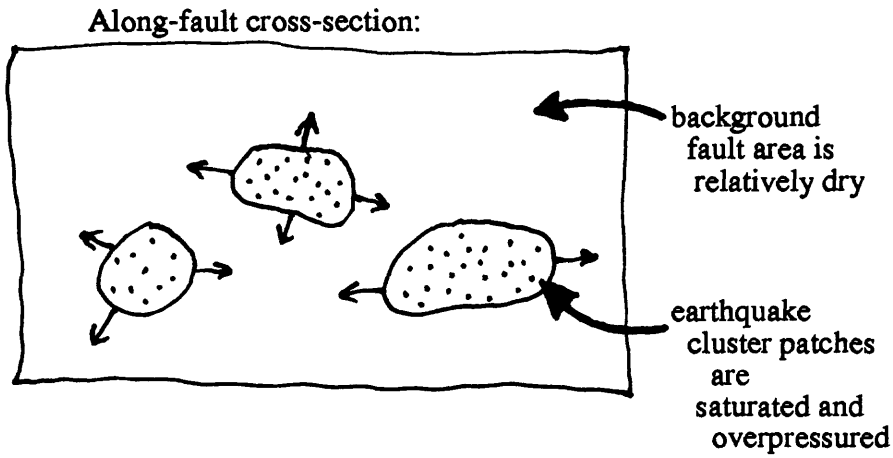



**Figure 4.** Map views of the high-resolution locations for two earthquake clusters. Each map is 500 m x 500 m; grid lines are 50 m apart. Map is oriented with respect to the San Andreas fault, which is parallel to the $y$-axis. The origin of the coordinate system is with respect to the 1966 epicenter: $y$-axis is distance in kilometers northwest from the 1966 epicenter; $x$-axis is distance in kilometers northeast from the 1966 epicenter. The *en echelon* trends discussed in the text are indicated by the orientation of the arrows just outside the map borders.

Model 1. Earthquake clusters occur in weak (saturated and overpressured) patches and the surrounding fault zone is strong (relatively dry).

Implications:
1. outward propagation with time of the earthquakes of a cluster
2. hydrofracturing




Model 2. Earthquake clusters occur in strong (relatively dry) patches and and the surrounding fault zone is weak (saturated and overpressured).

Implications:
1. inward propagation with time of the earthquakes of a cluster
2. conventional shear mechanisms

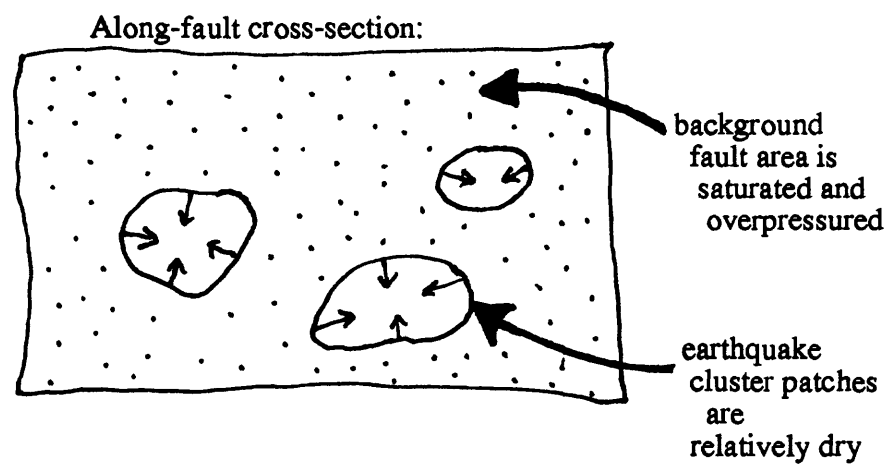



**Figure 5.** Models of fluid involvement in clusters; fluid-dominated (high pore pressure) mechanisms only (see text).

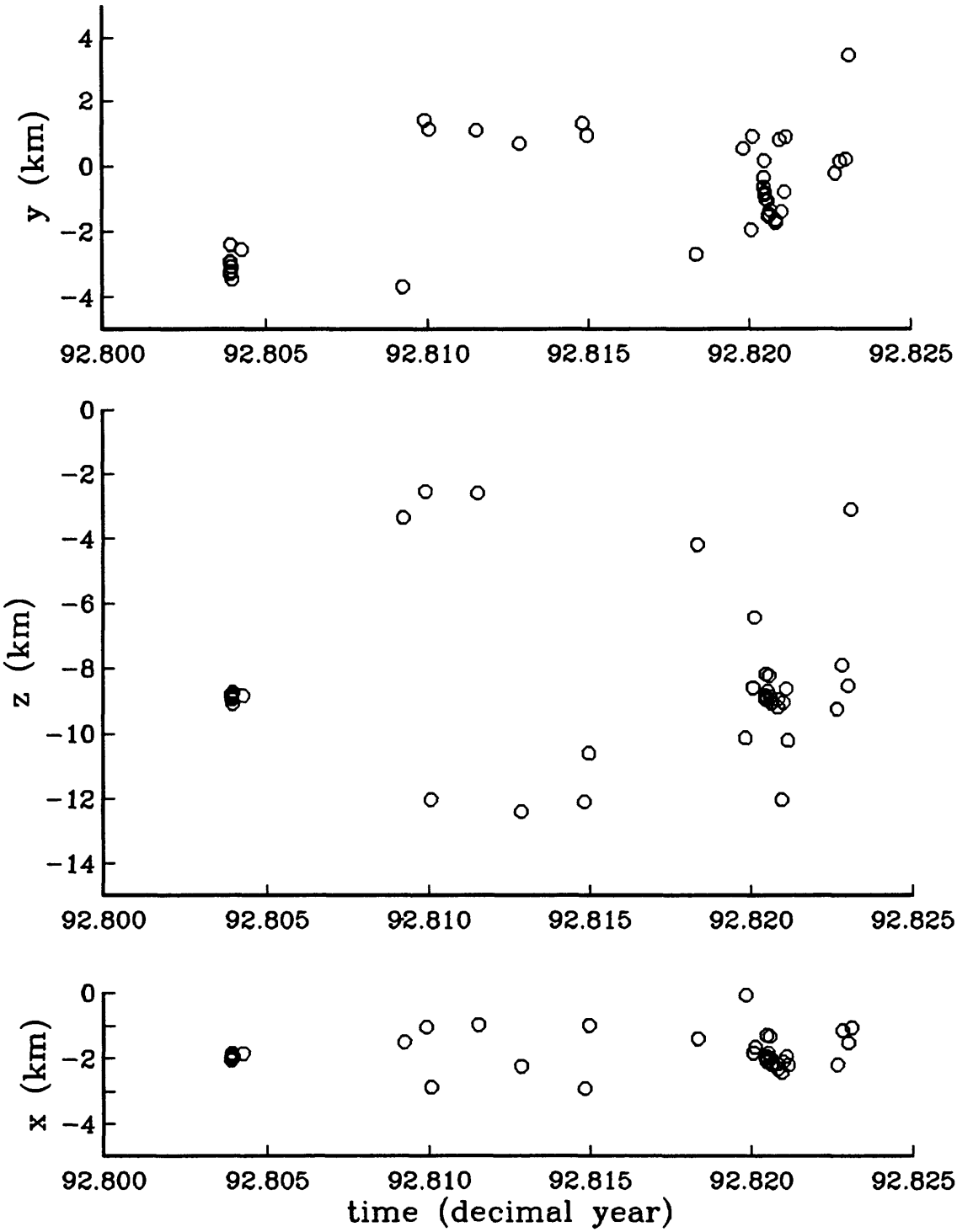



**Figure 6.** Spatial and temporal evolution of the earthquakes of a cluster which initiates with the M4.7 mainshock of October 1992 and spans a duration of approximately 8 days. Plots are $y$ versus time, $x$ versus time, and $z$ versus time, where $y$ is distance in kilometers northwest of the 1966 epicenter, $x$ is distance in kilometers northeast of the 1966 epicenter, and $z$ is depth. See text for discussion.

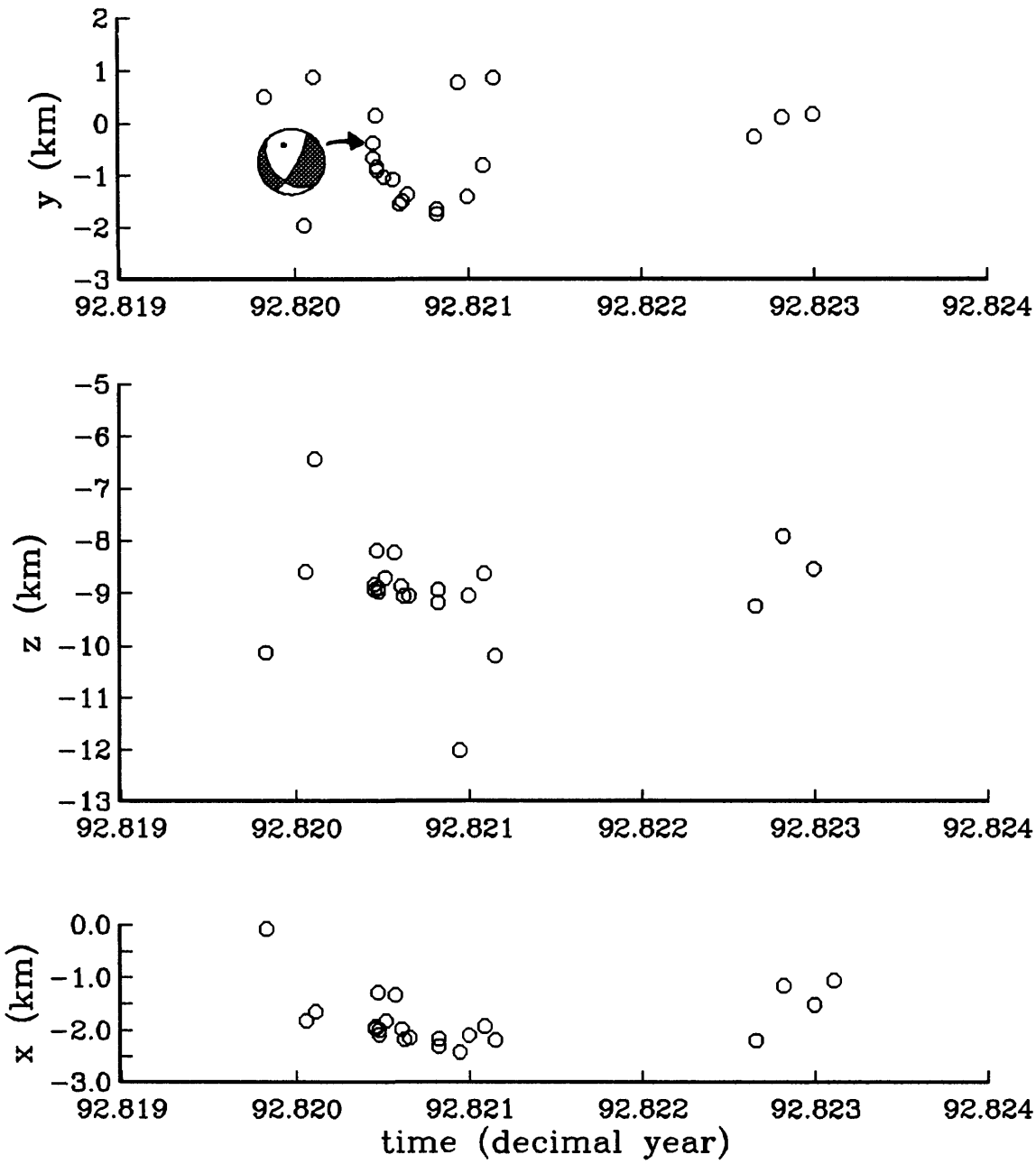



**Figure 7.** Spatial and temporal evolution of the last aftershocks in the October 1992 cluster shown in Figure 1. Plots are $y$ versus time, $x$ versus time, and $z$ versus time, where $y$ is distance in kilometers northwest of the 1966 epicenter, $x$ is distance in kilometers northeast of the 1966 epicenter, and $z$ is depth. Focal mechanism is for the M1.5 initiating event of the tight cluster starting at $t \approx 92.8204$, from the U.S. Geological Survey's catalog. See text for discussion.

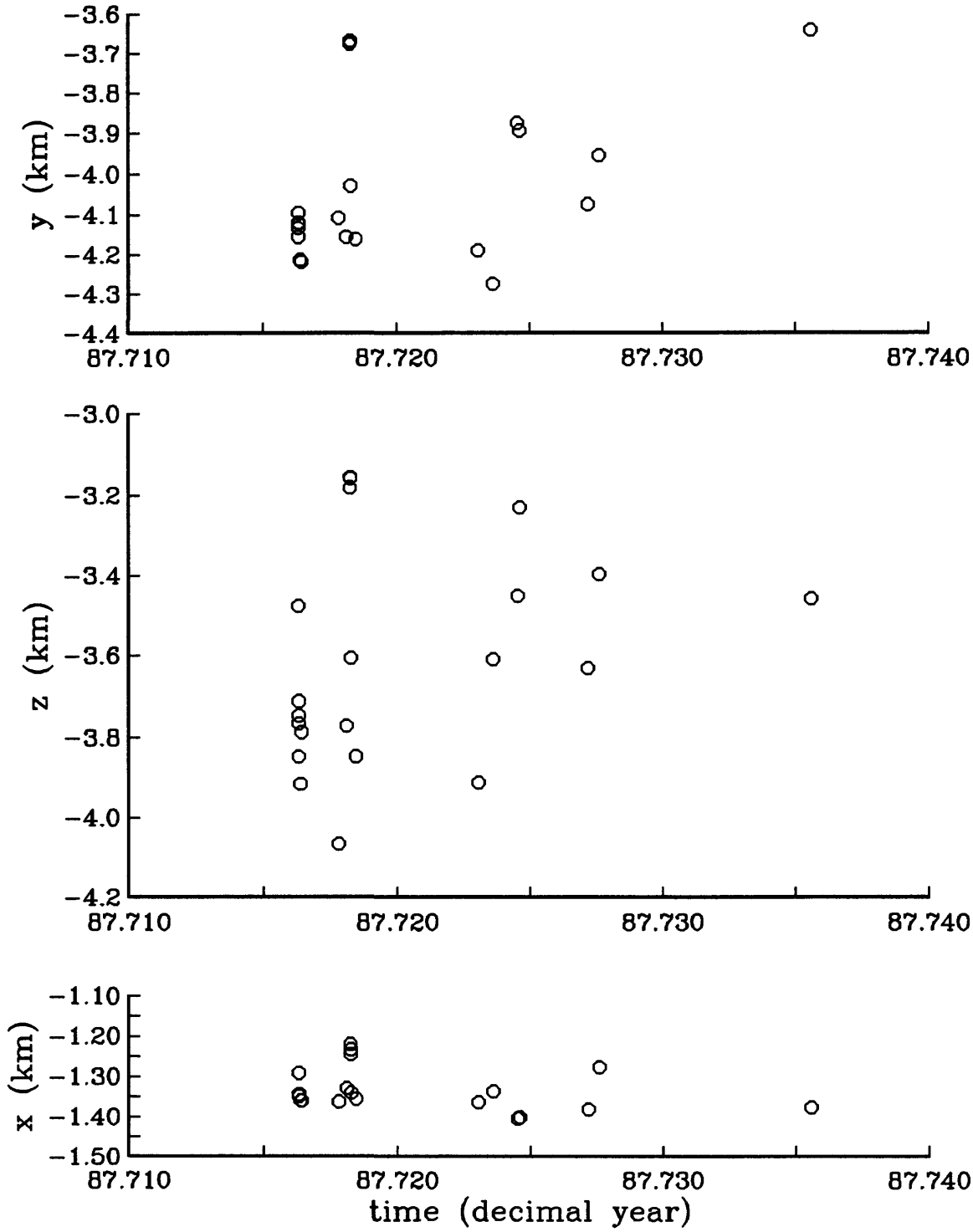



**Figure 8.** Spatial and temporal evolution of a cluster that occurred in 1987. Plots are $y$ versus time, $x$ versus time, and $z$ versus time, where $y$ is distance in kilometers northwest of the 1966 epicenter, $x$ is distance in kilometers northeast of the 1966 epicenter, and $z$ is depth. See text for discussion.

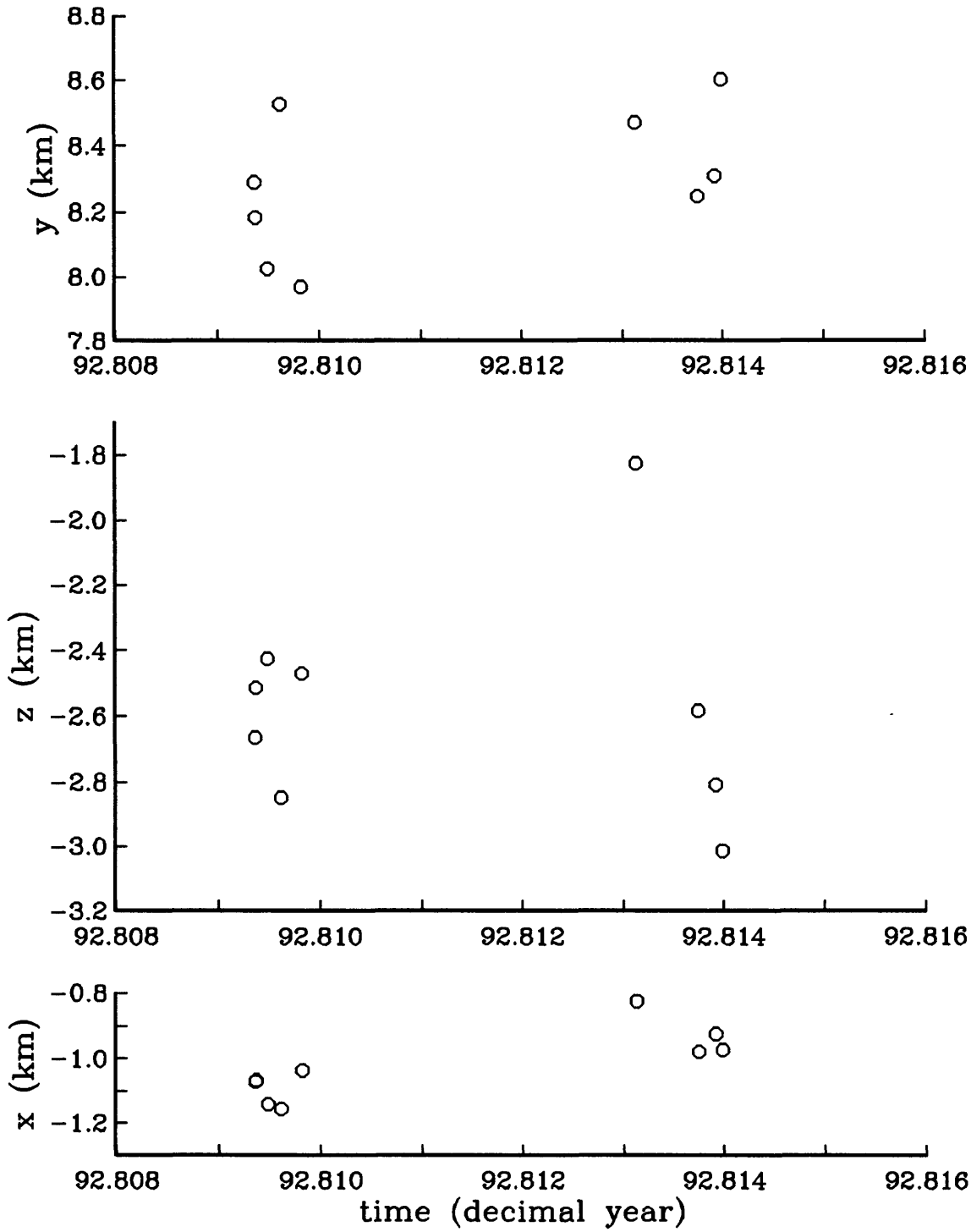



**Figure 9.** Spatial and temporal evolution of cluster that occurred during the October 1992 M4.7 sequence (see Figure 6) but farther northwest and at a shallower depth. Plots are $y$ versus time, $x$ versus time, and $z$ versus time, where $y$ is distance in kilometers northwest of the 1966 epicenter, $x$ is distance in kilometers northeast of the 1966 epicenter, and $z$ is depth. See text for discussion.

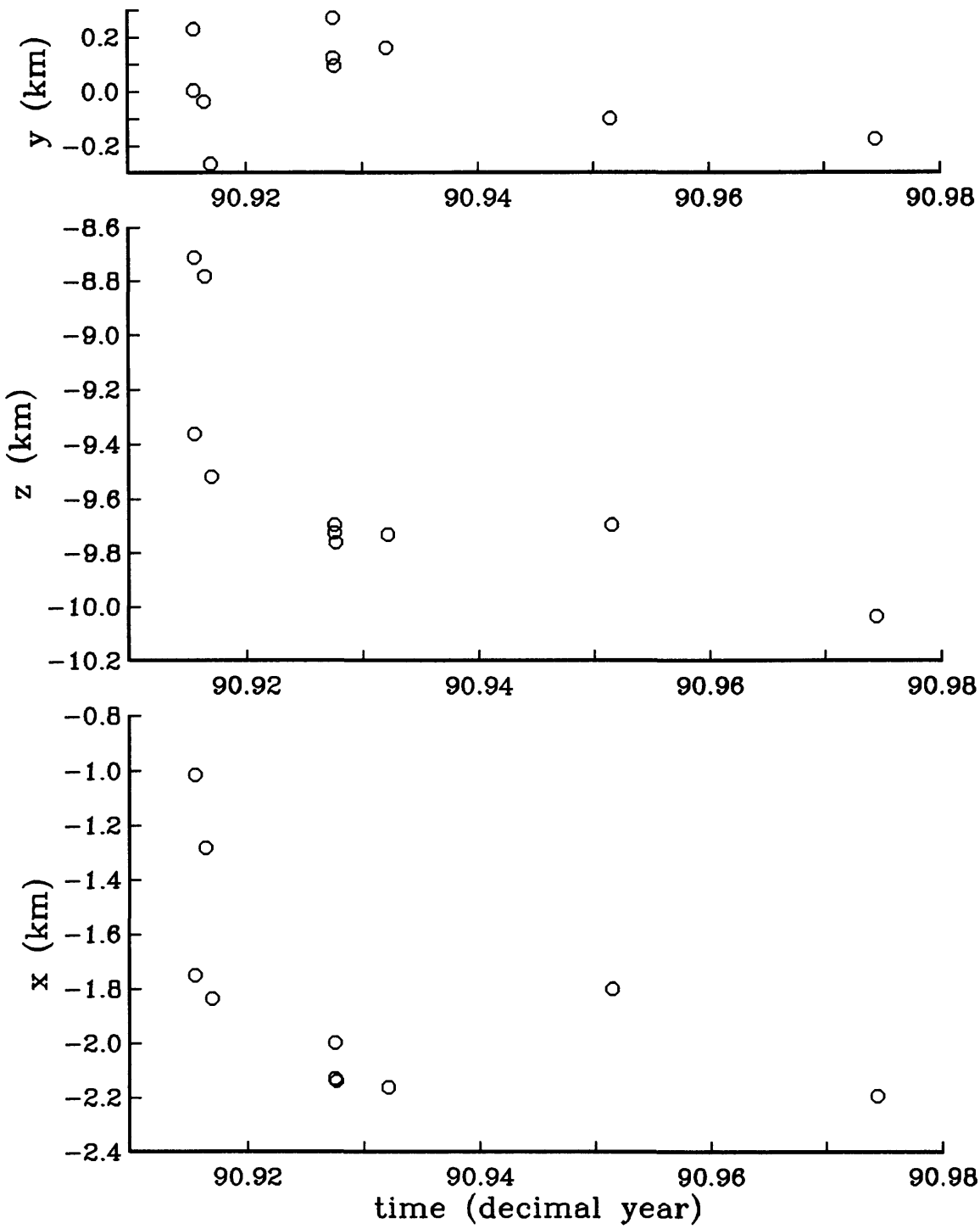



**Figure 10.** Spatial and temporal evolution of cluster that occurred in 1990 within a few kilometers of the October 1992 M4.7 sequence two years later (see Figure 6). Plots are $y$ versus time, $x$ versus time, and $z$ versus time, where $y$ is distance in kilometers northwest of the 1966 epicenter, $x$ is distance in kilometers northeast of the 1966 epicenter, and $z$ is depth. See text for discussion.

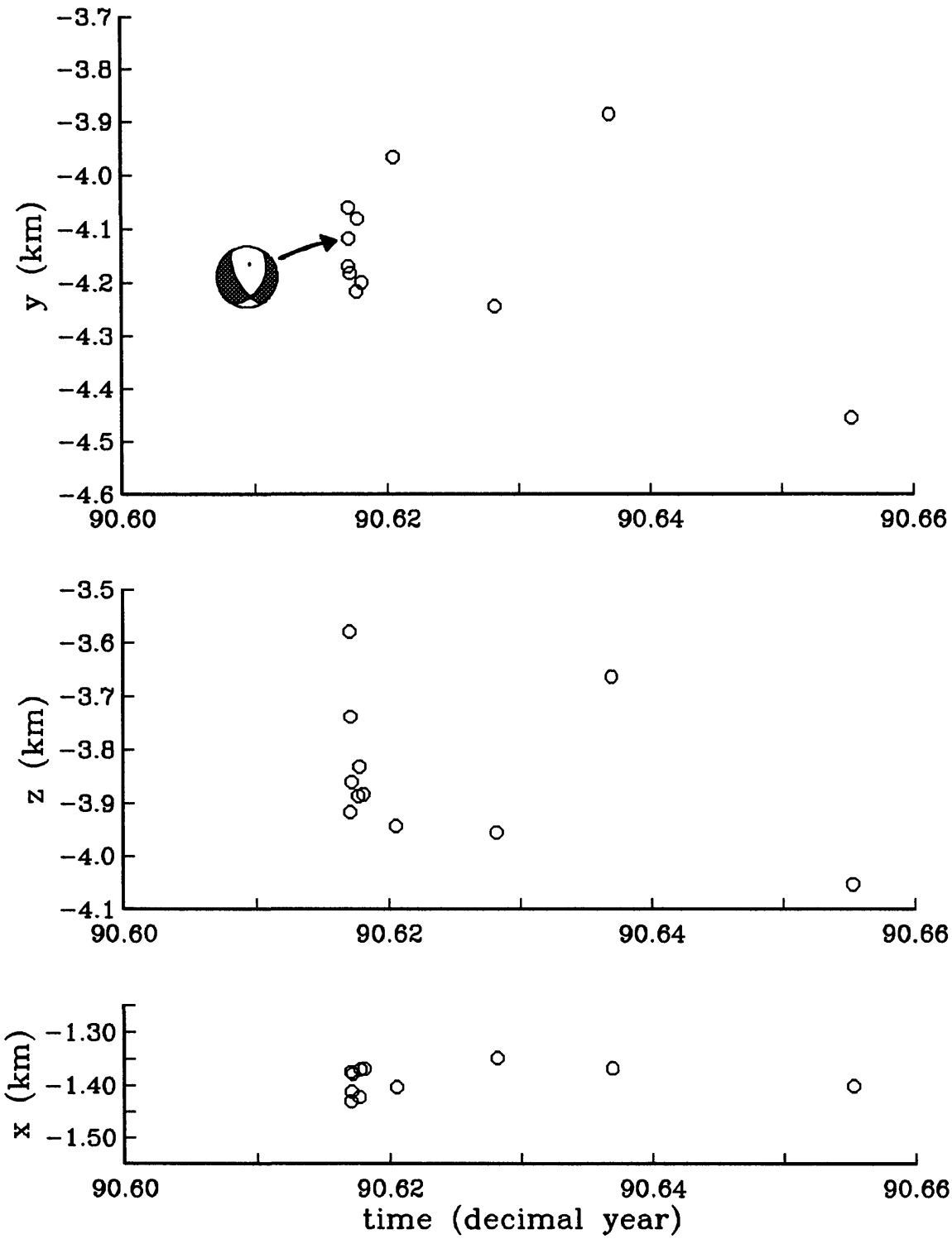



**Figure 11.** Spatial and temporal evolution of a cluster that occurred in 1990. Plots are $y$ versus time, $x$ versus time, and $z$ versus time, where $y$ is distance in kilometers northwest of the 1966 epicenter, $x$ is distance in kilometers northeast of the 1966 epicenter, and $z$ is depth. Focal mechanism (lower hemisphere) is for the M2.2 initiating event, from the U.S. Geological Survey's catalog. See text for discussion.

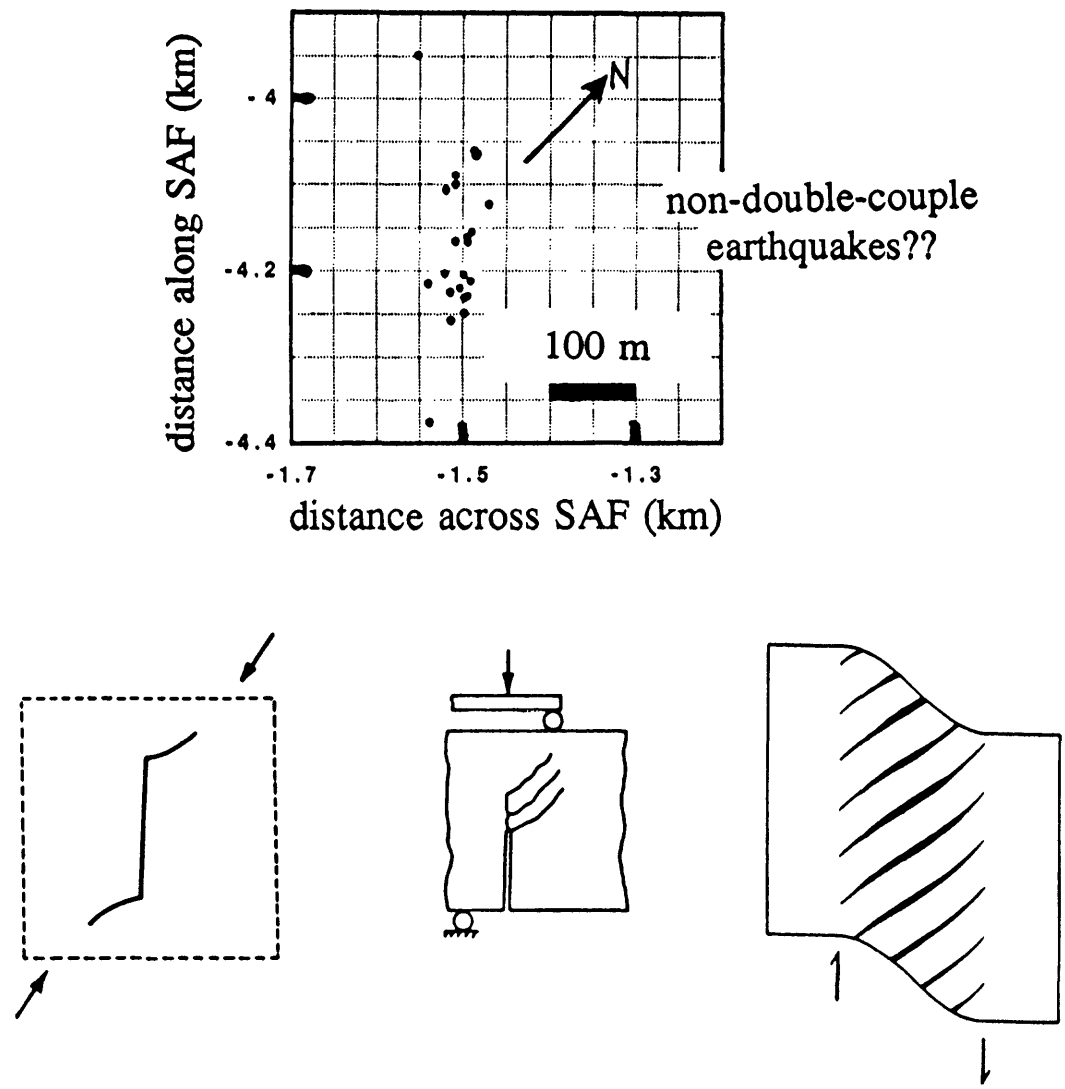



**Figure 12.** Analogy between the *en echelon* epicentral lineations characteristic of some Parkfield clusters (top) and coupled shear and tensile crack models (lower), including wing cracks (left and center; from Segall and Pollard, 1983) and an array of tension gashes (right; from Ramsay, 1967). The cluster shown is the same as shown in Figure 4a. In the map view of the cluster, the orientation of the San Andreas fault is parallel to the $y$-axis as are the shear cracks in the models below; the *en echelon* epicentral lineations are generally parallel to the wing cracks and tension gashes.

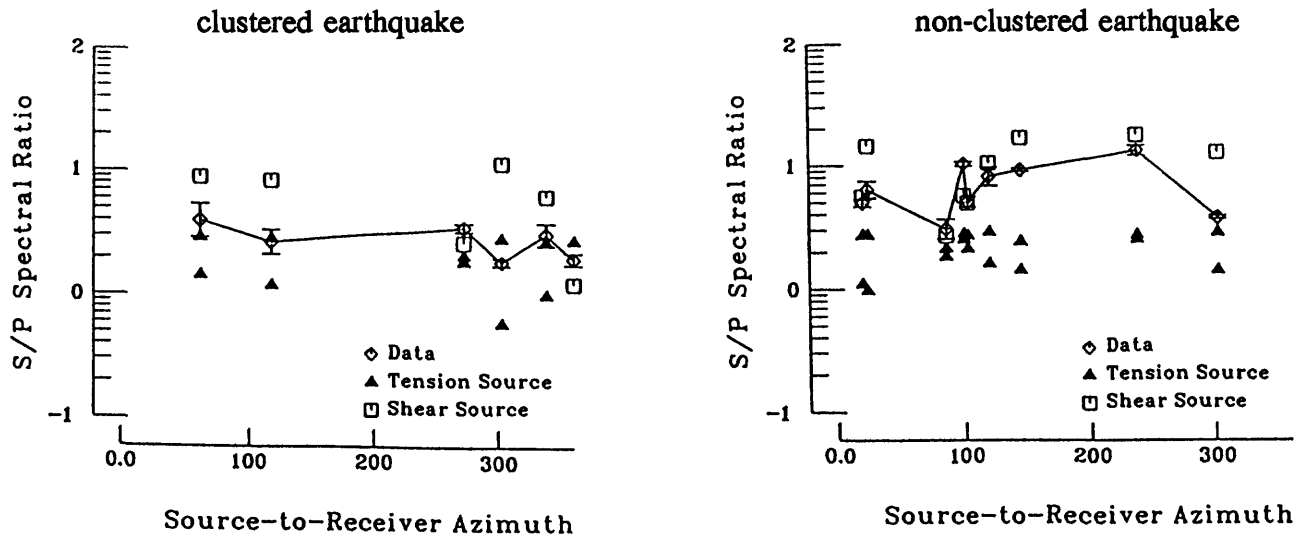



**Figure 13.** Log of S/P spectral ratios plotted versus source-to-receiver azimuth for a clustered event (left) and a non-clustered event (right). Lines joining diamonds connect the observed ratios at each site; standard errors are indicated. Also plotted are the calculated values of the S/P ratio for a shear source (open squares) and for two tensional sources (solid triangles). The expected shear source values are calculated using the catalog hypocentral locations, source incidence angles, epicenter-to-station azimuths, and the planes whose orientations agree with a pure-strike-slip source on the San Andreas fault (using its local azimuth of N45W), including its auxiliary plane. The double-couple gives a single calculated value at each azimuth. The predicted tensile values (solid triangles) are calculated for the correct orientation of a wing crack relative to its associated shear plane (70 degrees rotated from the slip vector clockwise or anticlockwise for right- or left-lateral slip). There are two sets of calculated tensile values because we have used the orientations of both the San Andreas fault and its auxiliary plane as the associated shear plane, which give different orientations of wing cracks. The ratios observed for the cluster event suggest a significant component of non-double-couple motion, while those observed for the non-clustered earthquake agree well with those predicted for a shear event on the San Andreas fault.

PKD1 930461711
Near-field Moment Tensor Solution
Mo = 4.0E+21 dyne-cm
Mw = 3.7
Weighted Correlation = 0.923
Delta = 14.3 km & Az = 143
Depth = 15.7 km & Half Duration = 0.262 sec
%DC = 78.2, %CLVD = 16.7 & %Iso = 5.1
P1: Strike, Rake & Dip = 316, -166 & 83
P2: Strike, Rake & Dip = 224, -7 & 76

PKD1 930940521
Near-field Moment Tensor Solution
Mo = 2.2E+23 dyne-cm
Mw = 4.9
Weighted Correlation = 0.962
Delta = 10.0 km & Az = 140
Depth = 8.2 km & Half Duration = 0.538 sec
%DC = 73.4, %CLVD = 14.9 & %Iso = 11.6
P1: Strike, Rake & Dip = 311, -175 & 85
P2: Strike, Rake & Dip = 220, -5 & 85

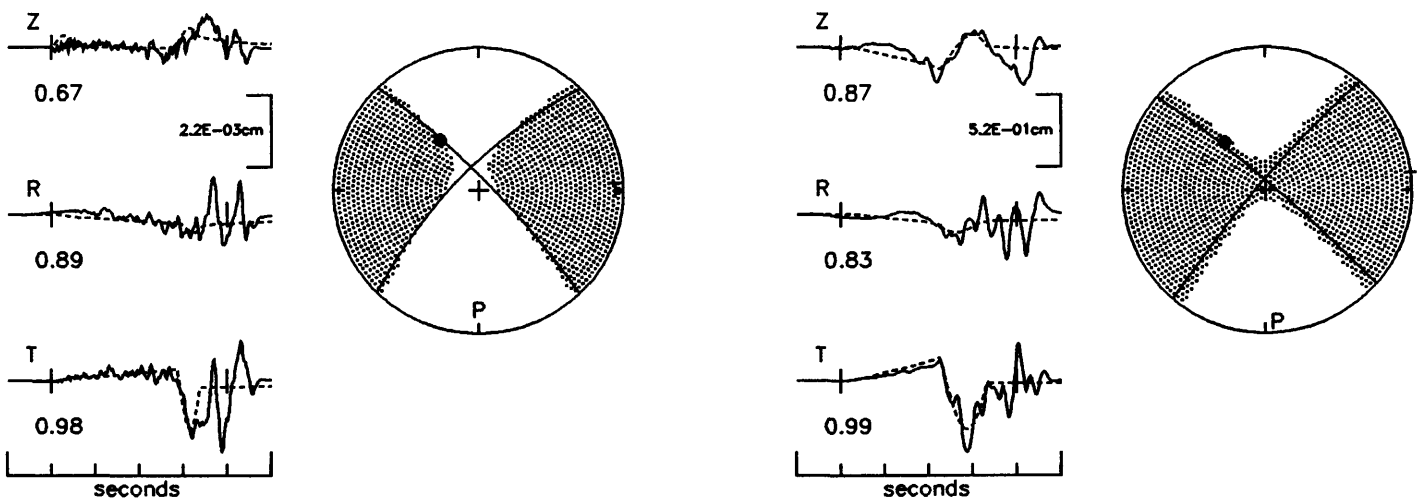



**Figure 14.** Near-field moment tensor inversions for two Parkfield events ($M_W$ 3.7, 15 February 1993; $M_W$ 4.9, 4 April 1993). Broadband data of UCB's PKD1 station were used (80 samples/second). Synthetic seismograms (dashed lines) were computed with a half-space model and are compared with the observed data (solid). The near-field ramps are visible on all three components (vertical, radial, and transverse). The orthogonal lines of the mechanisms represent the best-fitting double-couple solution; shaded regions represent the full unconstrained solution, which contains all six independent components of the moment tensor. Both events have a significant CLVD component (17% and 15%, respectively). Solutions by R. Uhrhammer.

# Triggered Earthquakes and Deep Well Activities[1]

CRAIG NICHOLSON[2] AND ROBERT L. WESSON[3]

[2]*Institute for Crustal Studies, University of California, Santa Barbara, California 93106-1100*

[3]*U.S. Geological Survey, MS905 National Center, Reston, Virginia 22092*

***ABSTRACT:*** Earthquakes can be triggered by any significant perturbation of the hydrologic regime. In areas where potentially active faults are already close to failure, the increased pore pressure resulting from fluid injection, or, alternatively, the massive extraction of fluid or gas, can induce sufficient stress and/or strain changes that, with time, can lead to sudden catastrophic failure in a major earthquake. Injection-induced earthquakes typically result from the reduction in frictional strength along pre-existing, nearby faults caused by the increased formation fluid pressure. Earthquakes associated with production appear to respond to more complex mechanisms of subsidence, crustal unloading, and poroelastic changes in response to applied strains induced by the massive withdrawal of subsurface material. As each of these different types of triggered events can occur up to several years after well activities have begun (or even several years after all well activities have stopped), this suggests that the actual triggering process may be a very complex combination of effects, particularly if both fluid extraction and injection have taken place locally. To date, more than thirty cases of earthquakes triggered by well activities can be documented throughout the United States and Canada. Based on these case histories, it is evident that, owing to preexisting stress conditions in the upper crust, certain areas tend to have higher probabilities of exhibiting such induced seismicity.

## INTRODUCTION

The phenomena of earthquakes triggered by deep well activities is certainly not new or unusual. *Richter* [1958] discusses the effects of shallow "slump earthquakes" within the Wilmington oil field, California, near the Los Angeles harbor in the years 1947, 1949, 1951 and 1955. Although he did not specifically correlate the earthquakes to the extensive ground subsidence caused by the massive withdrawal of oil and gas from the field, he did note their spatial coincidence, and mentioned a series of similar triggered "slump earthquakes" in the Po Valley, Italy, attributed by *Caloi et al.* [1956] to the commercial extraction of methane gas. In the 1920's, a series of "slight earthquakes" was felt near the Goose Creek oil field in south Texas, where oil production there had caused the field to subside by as much as 1 m between 1917 and 1925 [*Pratt and Johnson,* 1926; *Segall,* 1989]. Similarly, the injection of fluid at relatively high pressures can also induced adjacent seismicity, if the area is already close to failure. Most of these cases of seismicity related to fluid injection are associated with either water-flood operations to enhance the secondary recovery of hydrocarbons, or with the commercial stimulation (*i.e.,* hydraulic fracturing) of the well to increase fracture permeabilities [*Nicholson and Wesson,* 1990]. There are, however, a few specific cases in which waste disposal by fluid injection has also induced adjacent seismicity, including the largest and probably the best known earthquake to have been triggered by fluid injection—a magnitude 5.5 earthquake near Denver, Colorado in 1967 [*e.g., Healy et al.,* 1968].

In this paper, we survey a number of possible induced earthquakes related to adjacent deep well operations. Figure 1 shows locations within the United States and southern Canada where significant earthquakes have known to occur in close proximity to active well sites. In many cases, the only available evidence is simply a coincidence in space and time between specific earthquakes and known or inferred well activities. Few of these examples are well documented because of potential liability concerns of the respective well operators. We thus expect that many cases of possible induced deformation largely go unreported, either because the induced earthquakes are small (or the deformation is aseismic [*Davis and Pennington,* 1989]), or the well activities are not generally publicized.

---

[1] Reprinted from *Special Issue on Induced Seismicity,* A. McGarr, ed., *PAGEOPH,* **139,** n. 3-4, 1992.

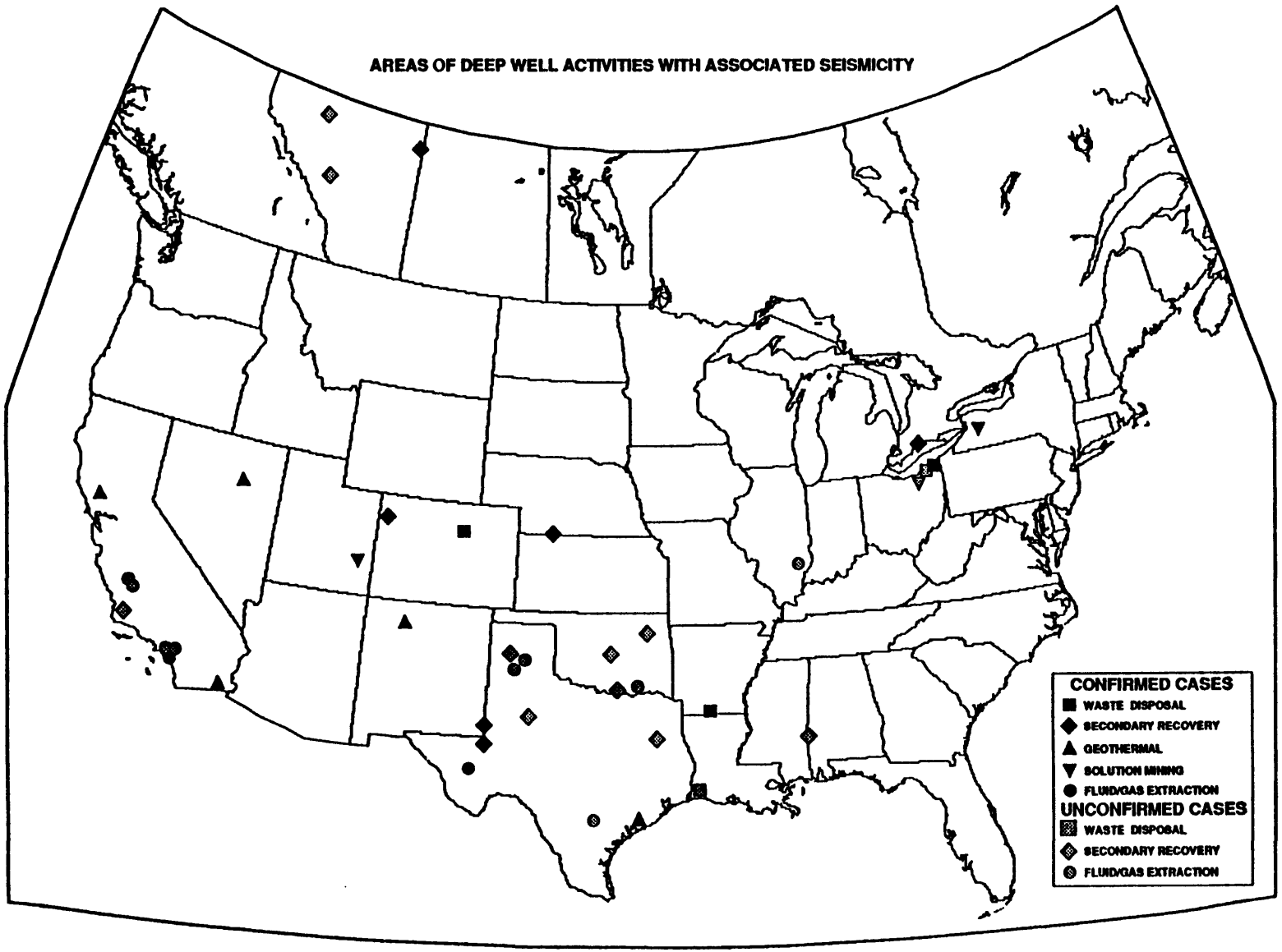



Figure 1. Sites of possible induced earthquakes that have occurred in close proximity to deep well activities. Symbol corresponds to predominate well operations at each site. Black symbols are sites that are either documented in the literature, or for which compelling evidence exists. Gray symbols are less well documented examples—for which, in many cases, only a simple spatial and temporal correlation exists.

## CASE HISTORIES

Earthquakes related to well activities typically fall into two major classifications. Those that are largely related to fluid injection and the resulting increased pore fluid pressure this may cause [*e.g., Healy et al.,* 1968; *Raleigh et al.,* 1976], or those that appear to have occurred in areas that have experience massive withdrawals of subsurface fluid or gas [*e.g., Kovach,* 1974; *Yerkes and Castle,* 1976; *Pennington et al.,* 1986; *Wetmiller,* 1986; *Segall,* 1989; *Doser et al.,* 1991; *McGarr,* 1991]. Table 1 lists a number of well sites that may have triggered adjacent seismicity. Detailed summaries of many of these case histories can be found in *Nicholson and Wesson* [1990].

Although we make a distinction between earthquakes associated with fluid injection versus earthquakes associated with fluid (or gas) withdrawal, in many cases where documentation exists, both deep well activities (extraction and injection) have taken place locally (Table 1). The exceptions tend to be either waste disposal, geothermal or stimulation operations (for fluid injection), or cases in which oil or gas production clearly predominate (for fluid extraction). Almost all other cases of earthquakes related to deep well activities seem to occur in areas where prior production decreased local formation pore pressure sufficiently to necessitate the initiation of secondary recovery operations. In such cases, the determination of the actual earthquake trigger process, or the mechanism of strain localization that lead to failure, is a much more difficult procedure. The fact that injection operations have taken place locally may be only incidental to the crustal readjustments or poroelastic strain changes that were triggered in response to the massive withdrawal of subsurface material [*e.g., Segall,* 1989]. Here we attempt to present case histories that are sufficiently distinct that some understanding of the dominant physical mechanism of the earthquake trigger process can be recognized.

### Overview of earthquakes induced by deep well fluid injection

Documented examples of seismic activity induced by fluid injection include earthquakes triggered by waste injection near Denver [*Healy et al.,* 1968] and in south-central Arkansas [*Cox,* 1991]; by secondary recovery of oil in Colorado [*Raleigh et al.,* 1972], southern Nebraska [*Rothe and Lui,* 1983], West Texas [*Harding,* 1981; *Davis,* 1985; *Davis and Pennington,* 1989], western Alberta [*Milne and Berry,* 1976], and southwestern Ontario [*Mereu et al.,* 1986]; by solution mining for salt in western New York [*Fletcher and Sykes,* 1977]; and by fluid stimulation to enhance geothermal energy extraction in New Mexico [*Pearson,* 1981]. In one specific case near Rangely, Colorado [*Raleigh et al.,* 1976], an experiment to control directly the behavior of large numbers of small earthquakes by manipulation of the fluid injection pressure was conducted successfully. Other cases of triggered seismicity, which were the result of either fluid injection or reservoir impoundment, were reviewed and discussed by *Simpson* [1986].

Of the well-documented cases of earthquakes related to fluid injection, most are associated with water-flood operations for the purpose of secondary recovery of hydrocarbons. This is because secondary recovery operations often entail large arrays of wells injecting fluids at high pressures into small confined reservoirs that have low permeabilities. Often, the producing field is a structural trap that may be defined by fault-controlled boundaries. In contrast, waste-disposal wells typically inject at lower pressures into large porous aquifers with high permeabilities that are away from known fault structures. This explains, in large part, why, of the many hazardous and nonhazardous waste-disposal wells in the United States, only three have ever been conclusively shown to be associated with triggering significant adjacent seismicity. These are wells located near Ashtabula, Ohio, El Dorado, Arkansas, and Denver, Colorado.

In the case near Ashtabula, a series of small shallow earthquakes was triggered close to the bottom of a 1.8-km deep well (RS#1, Figure 2); the largest of these was a magnitude 3.6 earthquake that occurred in 1987 [*Armbruster et al.,* 1987]. The injection well had been in operation only since 1986, and typically operated at injection pressures of about 100 bars. Investigations of earlier earthquake activity in adjacent Lake County (Figure 2, top), indicated that injection pressures of 100 bars was more than sufficient to cause failure along favorably-oriented faults with frictional coefficients of 0.6 and cohesive strengths of 40 bars [*Nicholson et al.,* 1988]. As a result, waste disposal wells located near Perry, Ohio (CH#1 and CH#2, Figure 2, top),

Table 1. Possible and probable induced earthquakes in Canada and the United States associated with well operations.

| Well site or field location | Type of well operation | BHP* (bar) | Year production began–ended‡ | ΔBHP† (bar) | Year injection began–ended | Max $M_L$ | Year of quakes |
|---|---|---|---|---|---|---|---|
| Apollo-Hendrick, TX | production/second. recov. | ? | 1926- | ? | 1973- | 2.0 | 1978-79 |
| Ashtabula, OH | waste disposal | 191 | N/A | +100 | 1986- | 3.6 | 1987 |
| Attica, NY?? | solution salt mining | 40? | 1880's- | +50? | 1880's- | 5.2 | 1929-67 |
| Beowawe, NV | acid treatment/stimulation | ? | ? | +100 | 1983 | <-1 | 1983 |
| Coalinga, CA?? | oil production | 230 | 1905- | -120 | 1952, 1961 | 6.5 | 1983-85 |
| Cane Creek, SE Utah | solution potash mining | ? | 1964- | ? | 1970- | ~3.0 | 1980-84 |
| Catoosa, OK | gas withdrawal? | ? | 1941- | ? | ? | 4.7 | 1956-60 |
| Chaveroo field, NM | secondary recovery | ? | 1967?- | ? | 1992- | ? | 1992- |
| Cleveland, OH? | solution salt mining | ? | 1889- | ? | 1889- | 3.0? | 1898-07 |
| Cogdell field, TX | secondary recovery | 215 | 1949- | -136, +217 | 1953, 1956- | 4.6 | 1974-79 |
| Cold Lake, Alberta | second.recov./waste displ. | ? | ? | ? | ? | ~2.0 | 1984- |
| Dale, NY | solution salt mining | 43 | ? | +55 | 1971?- | 1.0 | 1971 |
| Denver (RMA), CO | waste disposal | 269 | N/A | +72 | 1962–1966 | 5.5 | 1962-67 |
| Dollarhide, TX-NM | secondary recovery | 228 | 1945-, 1949- | -165, +135 | 1959- | 3.5 | 1979?- |
| Dora Roberts, TX | secondary recovery | 324 | 1955- | -48, +431 | 1961- | 3.0 | 1964- |
| East Durant, OK | gas withdrawal | ? | 1958- | ? | ? | 3.5 | 1968 |
| East Texas, TX | production/second. recov. | 70 | 1930- | -103, +83 | 1942- | 4.3 | 1957 |
| El Dorado, AR | waste disposal | 230? | early-1920's- | +60 | 1970-, 1983- | 3.0 | 1983-91 |
| El Reno, OK?? | oil/gas withdrawal? | ? | 1910's- | ? | ? | 5.2 | 1918-79 |
| Fenton Hill, NM | geothermal/stimulation | 265 | N/A | +200 | 1979 | <1.0 | 1979 |
| Fashing field, TX | gas withdrawal | 352 | 1958- | -281 | N/A | 3.4 | 1973-83 |
| The Geysers, CA | geothermal | 35 | 1966- | -18 | 1966- | 4.0 | 1975- |
| Gobles field, Ontario | secondary recovery | 45 | 1960- | ? | 1969- | 2.8 | 1979-84 |
| Goose Creek, TX | oil production | ? | 1917-1925 | ? | ? | ? | 1920's |
| Hunt field, MS? | secondary recovery? | ? | ? | ? | ? | 3.6 | 1976-78 |
| Imogene field, TX | oil production | 246 | 1944- | -100 | N/A | 3.9 | 1973-83 |
| Inglewood field, CA? | production/second. recov. | 48? | 1924- | -39, >+100 | 1954, 1957- | 3.7 | 1962?- |
| Kettleman Hills, CA? | oil production | 325 | 1928-, 1930- | ? | 1967- | 6.1 | 1985 |
| Kermit field, TX | secondary recovery | 198 | 1950- | -185, +221 | 1958-, 1964- | 4.4 | 1964- |
| Keystone I&II, TX | secondary recovery | 204 | 1930-, 1943- | -100, +176 | 1962-, 1964- | 3.5 | 1964?- |
| Lake Charles, LA? | waste disposal? | ? | ? | +93 | ? | 3.8 | 1983- |
| Lambert field, TX | secondary recovery | 145 | 1979- | -15, +21 | 1979- | 3.4 | 1983-84 |
| Love Co., OK | second. recov./stimulation | ? | 1953- | -?, +277 | 1965-,1978- | 2.8? | 1977-79 |
| Monahans, TX | secondary recovery | 131 | 1961- | -185, +207 | 1965- | 3.0 | 1965- |
| Montebello, CA?? | oil production | ? | 1917-, 1938- | ? | 1953- | 5.9 | 1987 |
| Orcutt field, CA | stimulation/oil production | ? | 1901-, 1905- | -?, +183 | 1951-, 1991 | 3.5 | 1991 |
| Panhandle, TX | oil production | ? | 1910's- | ? | ? | 5.1 | 1925-66 |
| Paradise Valley, CO | brine disposal | ? | N/A | ? | late-1991 | 0.8 | late1991 |
| Perry, OH | waste disposal | 200 | N/A | +114 | 1975- | 2.7 | 1983-87 |
| Pleasant Bayou, TX | brine injection/geothermal | ? | 1962 | ? | 1962, 65, 70 | <1.5 | 1978-80 |
| Rangely, CO | second. recovery/research | 170 | 1945- | -?, +120 | 1957- | 3.1 | 1962-75 |
| Richland Co., IL?? | oil production | ? | ? | ? | 1952?- | 4.9 | 1987 |
| Salton Sea, CA | geothermal | ? | 1980's- | ? | 1988 | <2.0 | 1987-88 |
| Sleepy Hollow, NE | secondary recovery | 115 | ? | -?, +56 | 1966- | 2.9 | 1977-84 |
| Snipe Lake, Alberta | secondary recovery | ? | 1954- | ? | 1963- | 5.1 | 1970 |
| Strachan, Alberta | gas withdrawal | 500? | 1973- | -250 | N/A | 4.0 | 1974- |
| Tomahawk field, NM | brine disposal | ? | 1972?- | ? | ? | ? | ? |
| Ward-Estes, TX | secondary recovery | 103 | ? | -?, +117 | 1961- | 3.5 | 1964?- |
| Ward-South, TX | secondary recovery | 76 | ? | -?, +138 | 1960- | 3.0 | 1964?- |
| War-Wink, TX | gas withdrawal | 800+ | 1965- | ? | 1967-, 1969- | 3.0 | 1975-79 |
| Wilmington, CA | oil production | ? | 1932-, 1937- | ? | 1953- | 5.1 | 1947-61 |

* BHP – Bottom Hole Pressure – initial formation fluid pressure (in bars) at time of production or injection.

† ΔBHP – change in Bottom Hole Pressure relative to initial formation fluid pressure – maximum increase (+) or decrease (-) (in bars) during either injection (+) or extraction (-) operations.

‡ Multiple start years indicate dates when significate new production (or injection) began.

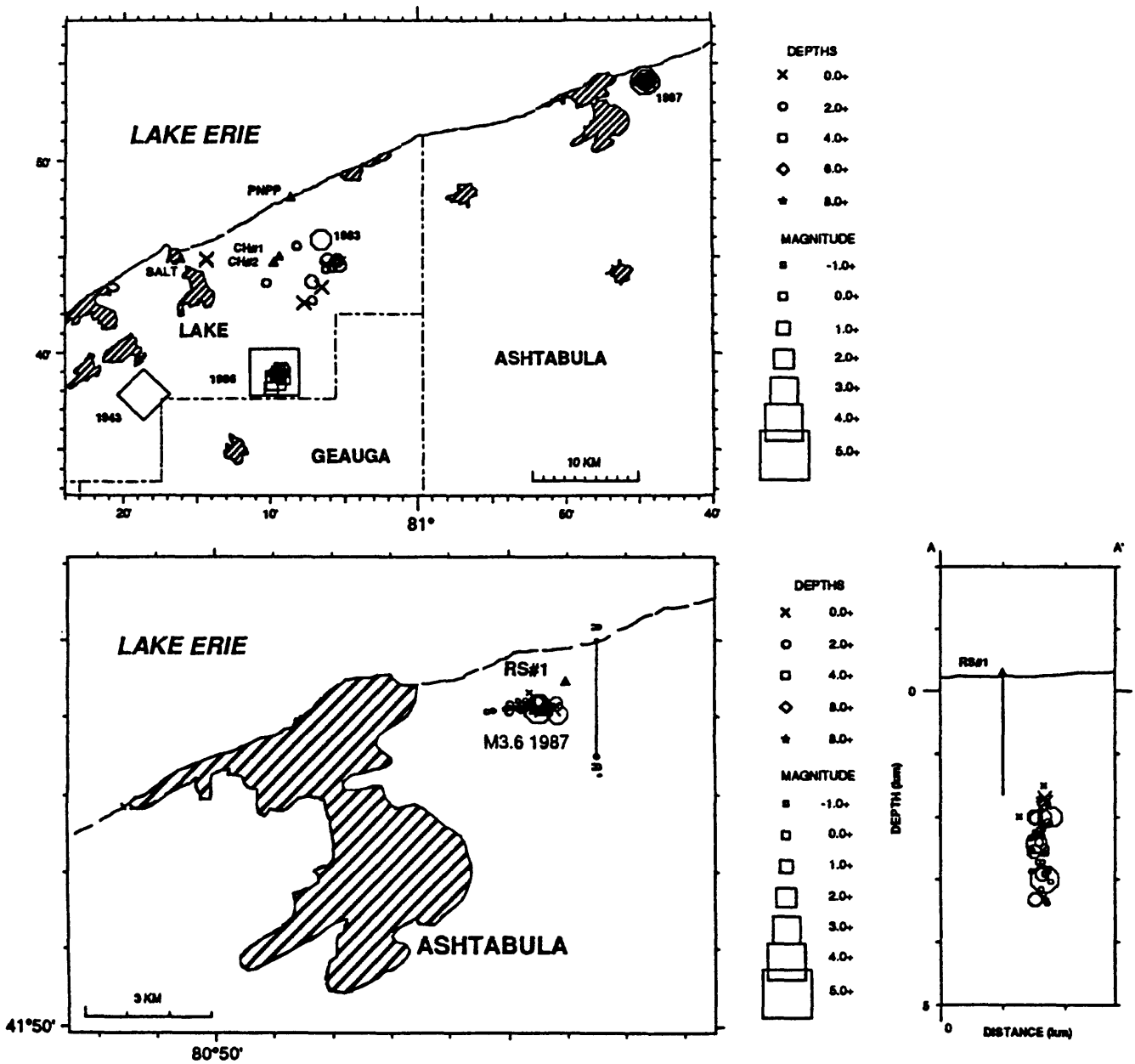



Figure 2. (top) Location of the 1987 induced earthquake sequence in northeastern Ohio near Ashtabula relative to earlier earthquakes (1943–1986) in Lake County. Striped areas are regions of dense population. Triangles are waste disposal wells; PNPP is the Perry Nuclear Power Plant (see *Nicholson et al.* [1988] for more details). (bottom) Map and cross section of the 1987 Ashtabula earthquake hypocenters relative to the location of a nearby active, high-pressure, waste-disposal injection well (Triangle, RS#1). Ashtabula earthquake data provided courtesy of John Armbruster.

apparently triggered several small earthquakes at distances less than 5 km. These wells operated at injection pressures of 100 bars or more. Other earthquakes located at greater distances (that included a magnitude 5.0 event in 1986) could not be sufficiently distinguished from natural background seismicity (such as the 1943 event) that their occurrence could be considered induced (Figure 2, top)[*Nicholson et al.,* 1988]; although earthquakes in southwestern Ontario [*Mereu et al.,* 1986] and western New York [*Fletcher and Sykes,* 1977] have been associated with similar adjacent deep well activities (Figure 1).

Near El Dorado, the disposal of waste brine under pressure, triggered a series of small earthquakes that have continued to the present; the largest of which was a magnitude 3.0 event in 1983. Although oil production had occurred in the area since the early 1920's and disposal of waste brine had begun in the 1970's, no earthquake activity was reported until 1983, shortly after large-volume injection activities began at relatively high pressures [*Cox,* 1991]. Since 1983, a strong correlation between rates of seismicity and sudden increases in injection volumes was also observed.

In the most prominent case of induced seismicity by fluid injection, the injection well responsible was located at the Rocky Mountain Arsenal near Denver, where fluid was being injected into relatively impermeable crystalline basement rock. This caused the largest known injection-induced earthquakes to date (three earthquakes between magnitude 5 and 5.5), the largest of which caused an estimated $0.5 million in damages in 1967. Although these induced earthquakes were by no means devastating, they did occasion extensive attention and concern and led, at least in the Denver case, to the cessation of all related injection well operations.

The Rocky Mountain Arsenal case is thus considered to be the classic example of earthquakes induced by deep well injection. Before this episode, the seismic hazard associated with deep well injection had not been fully appreciated. At the Rocky Mountain Arsenal, injection into the 3,700-m-deep disposal well began in 1962 and was quickly followed by a series of small earthquakes, many of which were felt in the greater Denver area (Figure 3). It was not until 1966, however, that a correlation was noticed between the frequency of earthquakes and the volume of fluid injected (Figure 4) [*Evans,* 1966]. Pumping ceased in late 1966 specifically because of the possible hazard associated with the induced earthquakes; after which, earthquakes near the bottom of the well stopped. Over the next 2 years, however, earthquakes continued to occur up to 6 km away from the well as the anomalous pressure front, which had been established around the well during injection, continued to migrate outward from the injection point (Figure 3). The largest earthquakes in the sequence (with magnitudes between 5.0 and 5.5) occurred in 1967, long after injection had stopped and well away from the point of fluid injection itself.

These results imply that the fluid pressure effects from injection operations can extend well beyond the expected range of actual fluid migration. Indications have shown, however, that the risk posed by such triggered earthquakes can be mitigated by careful control of the activity responsible for the induced seismicity. As shown by a number of cases [*Nicholson and Wesson,* 1990], seismicity eventually can be stopped either by ceasing the injection or by lowering pumping pressures. The occurrence of the largest earthquakes involved in the Rocky Mountain Arsenal case a year after all pumping had ceased, however, indicates that the process, once started, may not be controlled completely or easily.

In each of the well-documented examples of earthquakes associated with deep injection wells, convincing arguments that the earthquakes were induced relied upon three principal characteristics of the earthquake activity. First, there was a very close geographic association between the zone of fluid injection and the locations of the earthquakes in the resulting sequence. Second, calculations based on the measured or the inferred state of stress in the Earth's crust and the measured injection pressure indicated that the theoretical threshold for frictional sliding along favorably oriented preexisting fractures likely was exceeded. And, third, a clear disparity was established between any previous natural seismicity and the subsequent earthquakes, with the induced seismicity often characterized by large numbers of small earthquakes at relatively shallow depths that persisted for as long as elevated pore pressures in the hypocentral region continued to exist.

Many of the sites where injection-induced earthquakes have occurred operate at injection pressures above 100 bars ambient (Table 1). The exceptions tend to be sites characterized by a

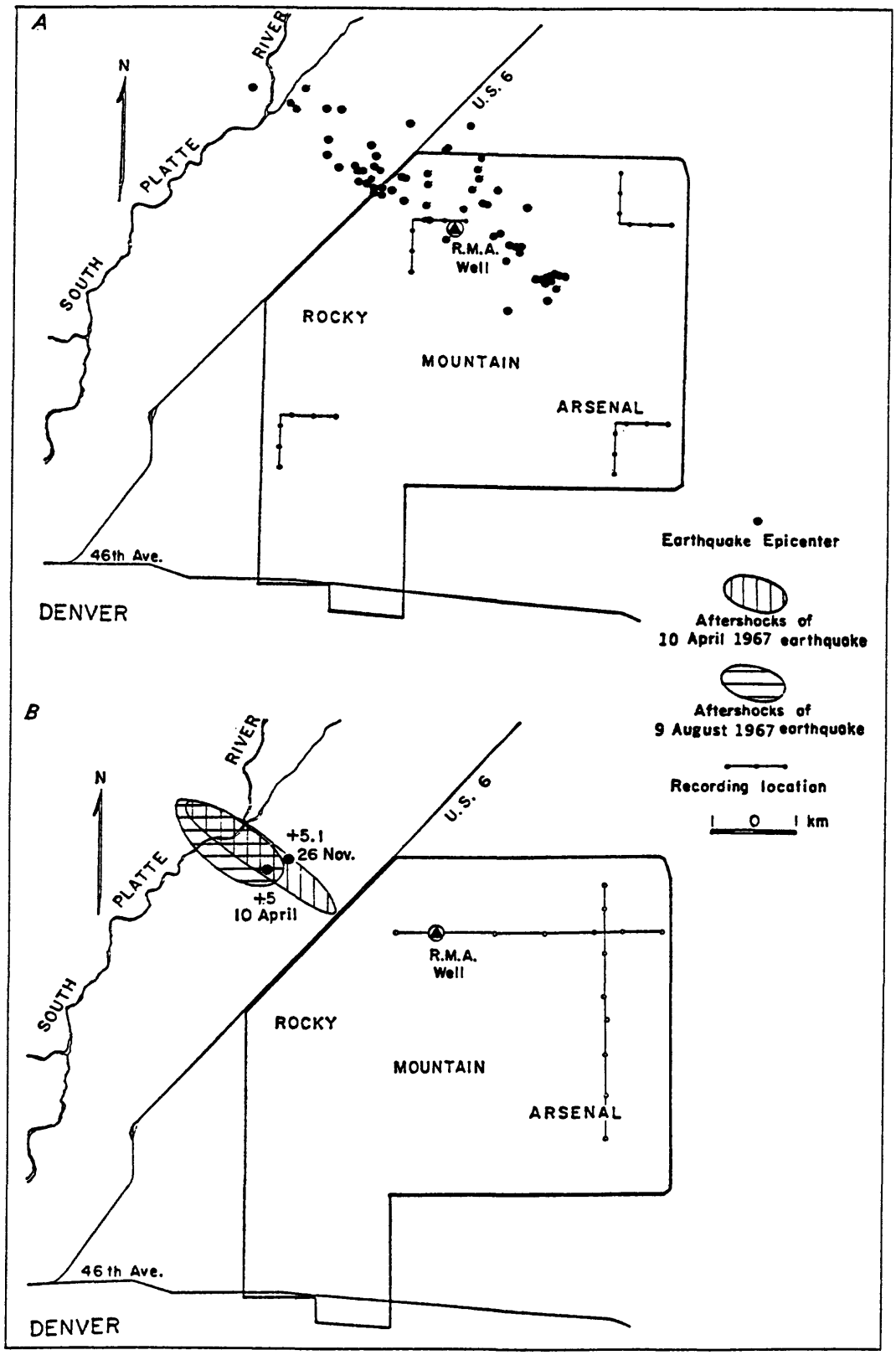



Figure 3. Earthquake activity near the Rocky Mountain Arsenal waste-disposal well, Colorado. (A) Epicentral distribution of earthquakes during January and February 1966. (B) Aftershock distributions of the large 1967 earthquakes. Reprinted with permission from *Healy et al.* [1968].

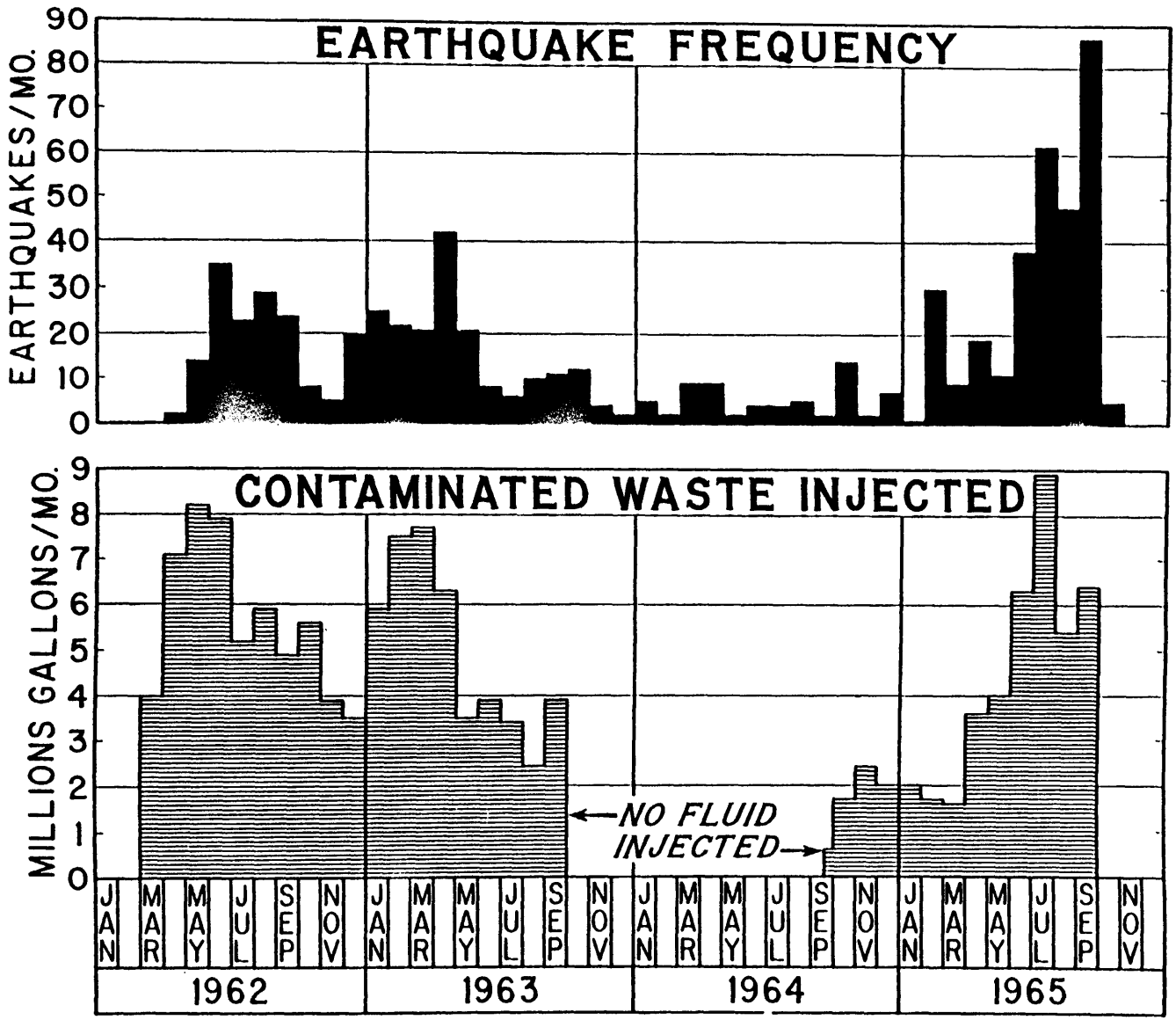



Figure 4. Correlation between earthquake frequency (top) and volume of contaminated waste injected (bottom) at the Rocky Mountain Arsenal well, Colorado. Reprinted with permission from *Healy et al.* [1968].

close proximity to recognized surface or subsurface faults. In the Rangely and the Sleepy Hollow Oil Field cases, faults are located within the pressurized reservoir and were identified on the basis of subsurface structure contours. At Attica and Dale, New York, the earthquakes occurred close to a prominent fault zone exposed at the surface (the Clarendon-Linden fault system). At the Rocky Mountain Arsenal well, fluid was inadvertently injected directly into a major subsurface fault structure, which was identified later only on the basis of the subsequent induced seismicity [*Healy et al.,* 1968] and the properties of the reservoir into which fluid was being injected, as reflected in the pressure-time record [*Hsieh and Bredehoeft,* 1981].

**Overview of earthquakes related to massive fluid or gas extraction**

Triggered earthquakes that spatially correlate with areas of massive fluid (or gas) withdrawal often fall into two distinct categories: (1) shallow induced earthquakes—within or near the producing formation—that typically exhibit normal or reverse faulting focal mechanisms and may be associated with the rapid subsidence and poroelastic strain changes resulting from the large volumes of material extracted [*Kovach,* 1974; *Yerkes and Castle,* 1976; *Wetmiller,* 1986; *Segall,* 1989; *Doser et al.,* 1991]; or (2) deep induced earthquakes—which may occur near the base of the seismogenic zone—that often exhibit thrust mechanisms and may be related to stress and/or strain changes associated with unloading effects caused by the large amounts of material locally removed from an area experiencing crustal convergence [*e.g., Simpson and Leith,* 1985; *McGarr,* 1991]. These latter induced events can be much larger in magnitude and can be much more difficult to distinguish than the shallow induced earthquakes, as the shallow seismicity is much more likely to exhibit temporal variations that correlate with specific activities at the adjacent producing wells.

One of the best examples of shallow induced earthquakes related to fluid withdrawal occurred near Los Angeles, California (Figure 5). The massive withdrawal of oil from one of the largest fields in the basin, the Wilmington oil field, resulted in significant subsidence within the city limits of Long Beach (Figure 6, top). Up to 8.8 m of surface subsidence was observed over an elliptically-shaped area between 1928 and 1970. This rapid subsidence, which reached a maximum rate of 71 cm/yr in 1951, 9 months after peak oil production, resulted in several damaging earthquakes, specifically in the years 1947, 1949, 1951, 1954, 1955, and 1961 (Figure 6, bottom) [*Kovach,* 1974]. In most cases, the earthquakes were unusually shallow and generated high intensities for their size. The largest earthquake occurred in 1949, and caused nearly 200 wells to go off production, many of them permanently [*Richter,* 1958]. Damage was estimated to be in excess of $9 million. The area affected equaled over 5.7 $km^2$ and involved measured displacements of 20 cm. This would correspond to an earthquake of moment magnitude 4.7, and is consistent with a magnitude of 5.1 estimated from the unusually well developed surface waves generated by the event [*Kovach,* 1974].

*Segall* [1989] presents a model that explains the local subsidence and the occurrence of shallow thrust faulting above and below the producing horizon as a result of poroelastic effects in response to reservoir compaction. The massive fluid extraction causes the reservoir rock to contract. This in turn induces the strata above and below the reservoir to be driven into local compression. Areas farther away from the reservoir are displaced less than the rock immediately above and below the reservoir—which causes flanking regions to extend and may result in normal faulting near the edges of the producing area [*Segall,* 1989]. Oil production from the Inglewood field (Figure 5) within the Los Angeles basin produced near-surface extensional creep events near the edge of the field starting in 1952 [*Hamilton and Meehan,* 1971]. Water-flood operations that began in 1954 apparently accelerated this normal faulting, produced increased shallow seismicity starting in 1962, and eventually lead to the failure of the Baldwin Hills water storage facility in 1963 that killed 5 people and caused $12 million damage.

In contrast, the earthquakes in the Wilmington oil field were apparently generated by low-angle reverse slip on bedding planes at depths of 470 to 520 m, while virtually all the compaction that caused the surface subsidence was localized in the producing beds at depths of 650 to 1050 m [*Kovach,* 1974; *Allen and Mayuga,* 1970; *Segall,* 1989]. Water flooding of the Wilmington field and adjacent areas was initiated in 1954 in an attempt to halt subsidence and to enhance the

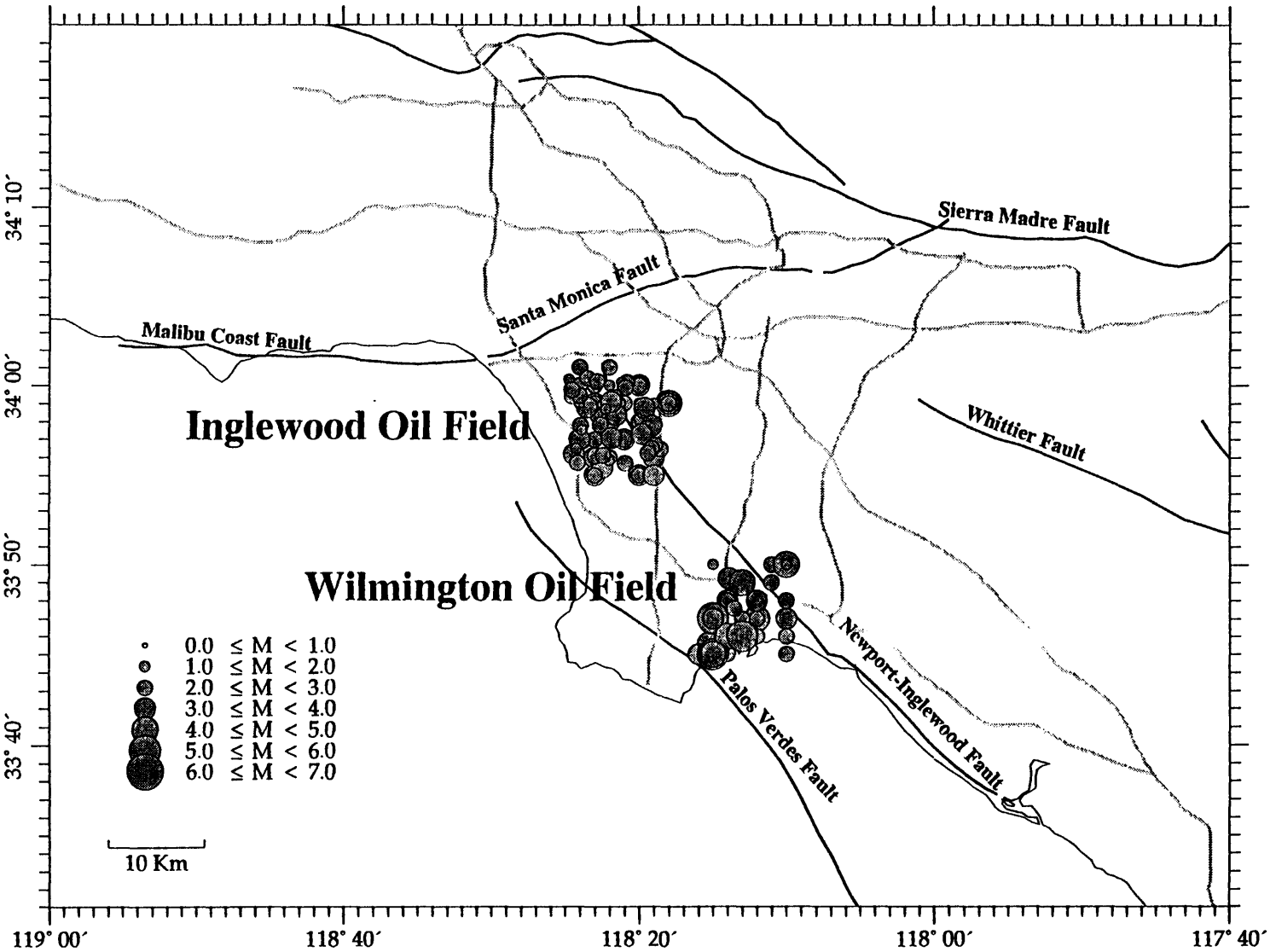
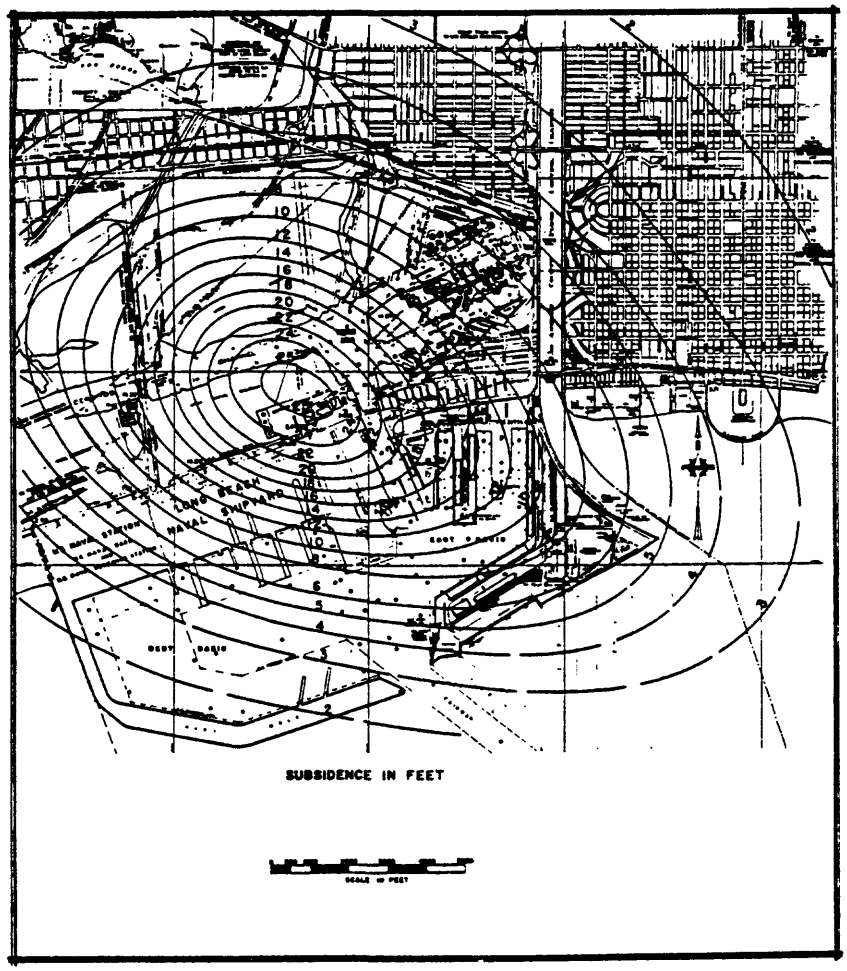



Figure 5. Map of the Los Angeles basin region showing possible subsidence-related earthquakes in the Wilmington (1937-1962) and Inglewood (1932-1976) oil fields.



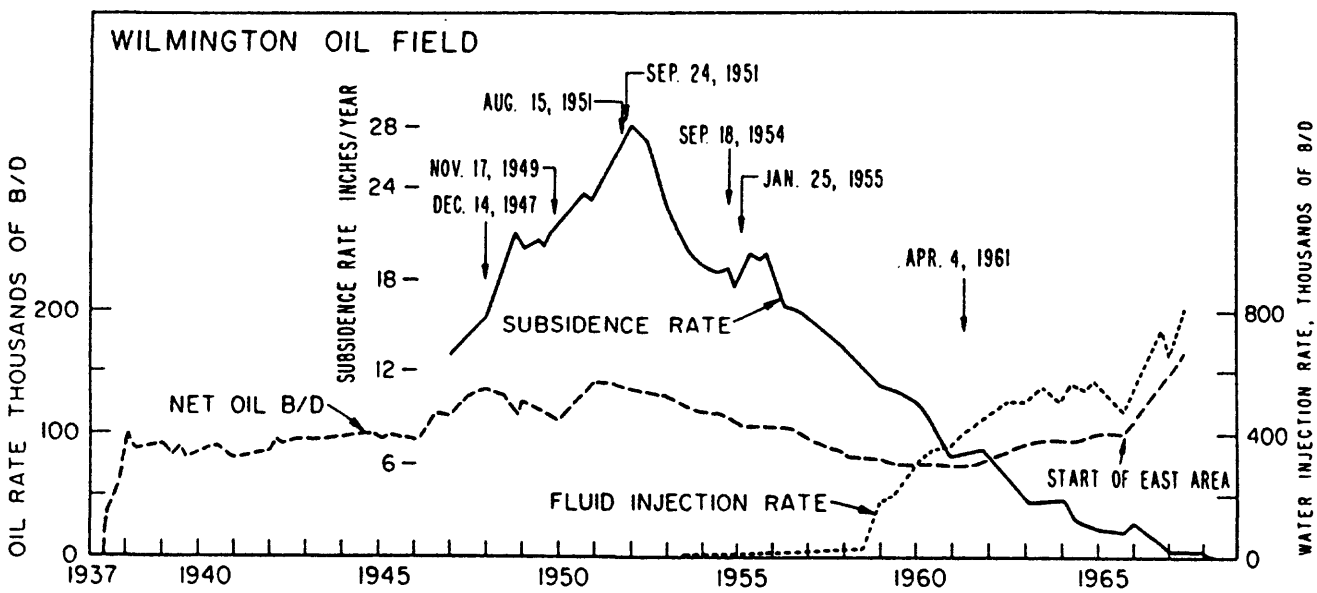




Figure 6. (top) Observed subsidence in feet for the year 1958 within the Wilmington oil field, California; data courtesy of the City of Long Beach. (bottom) Subsidence rate in the center of the Wilmington oil field compared with oil production and water injection rates. Arrows are dates of major damaging earthquakes. Reprinted with permission from *Kovach* [1974].

secondary recovery of oil. *Teng et al.* [1973] reported on the seismic activity associated with fourteen oil fields operating within the Los Angeles Basin where water-flood operations were taking place. Although much of the seismicity in the area is natural and occurs predominantly at depths as deep as 16 km along the Newport-Inglewood fault, seismic activity during 1971 appeared to correlate, at least in part, with injection volumes from nearby wells [*Teng et al.,* 1973]. However, many of the earthquakes detected were small (less than magnitude 3.0) and occurred at depths of 5 km or more, which made it difficult to distinguish them from the natural background seismicity. Subsequent injection operations have since stabilized to the point where fluid injection nearly equals fluid withdrawal and little, if any, seismic activity can be directly attributable to injection well operations.

Similar shallow induced earthquake activity has been associated with the withdrawal of natural oil and gas from several fields in Texas, including the Goose Creek, Fashing, Imogene and War-Wink fields [*e.g., Pratt and Johnson,* 1926; *Pennington et al.,* 1986; *Doser et al.,* 1991]. No significant injection operations had taken place in these areas, so much of the induced earthquake activity was inferred to be related to local subsidence and compaction caused by production [*Segall,* 1989]. The earthquakes associated with the Goose Creek field in south Texas occurred on normal faults that broke the surface along the northern and southern margins of the subsiding region [*Pratt and Johnson,* 1926]. Detailed analysis of the seismicity near the War-Wink field [*Doser et al.,* 1991] confirmed that many of the earthquakes were localized above and below producing horizons, and that many events exhibited normal and reverse focal mechanisms. However, not all the earthquake activity could be explained so easily or simply. Some of the earthquakes occurred at depths greater than 4 km. Poroelastic effects associated with reservoir compaction are limited to the shallow crust near the producing horizon [*Segall,* 1989]. The deep crustal earthquakes associated with areas of massive fluid extraction thus require further explanation. Possible mechanisms proposed to account for various aspects of the observed seismicity near the War-Wink field include: (1) the natural occurrence of deep (>4 km) crustal earthquakes along preexisting faults in response to the existing regional stress field, (2) strain localization in naturally over-pressured zones resulting from production, or (3) shallow earthquakes resulting from compaction or the upward migration of high fluid pressures [*Doser et al.,* 1991].

The possible correlation between very deep (>10 km) earthquake activity and the massive extraction of fluid or gas from the shallow crust is much more ambiguous [*McGarr,* 1991]. In such cases, the triggering mechanism is thought to be related to the isostatic imbalance caused by the net extraction of fluid or gas from the upper crust. As the ductile lower crust will deform in response to this imbalance, this readjustment will increase the applied load in the upper seismogenic layer, which may then fail seismically so as to restore local static equilibrium. If the upper crust has been already folded and faulted in response to applied horizontal tectonic compression, the stress and strain readjustments associated with restoring the isostatic imbalance will be concentrated on these preexisting structures. Thus, earthquakes near the base of the upper crust may be an expected outcome of major oil or gas production from growing anticlines—irrespective of the depths of the producing formation [*McGarr,* 1991].

Possible candidates for this particular type of triggered deep seismicity include the 1983 $M_L$ 6.5 Coalinga earthquake beneath the Coalinga oil field, the 1985 $M_L$ 6.1 Kettleman Hills earthquake beneath the Kettleman North Dome oil field, and the 1987 $M_L$ 5.9 Whittier Narrows earthquake beneath the Montebello oil field (Table 1). In each case, a mechanical connection is suggested between oil production and the earthquakes because—in each case—the total seismic deformation (moment) released during these earthquake sequences is nearly equal to that required to offset the force imbalance caused by the oil production [*McGarr,* 1991]. In addition, all three events exhibited nearly pure reverse motion on low-angle faults that core active folds responsible for originally trapping the oil.

An alternate explanation for this deep seismicity may be related to possible naturally occurring fluids in the deep crust. If, for example, preexisting high fluid pressures (approaching lithostatic) exist in the mid-to-lower crust, then the massive withdrawal of near-surface material may induce sufficient changes in the local hydrogeologic regime that significant migration of fluids at depth is

initiated. Thus, crustal unloading as a result of production may create pressure differentials that cause a preferential migration of deep fluids towards the producing area. The enhanced saturated condition beneath the producing area may then lend itself to increased strain localization, and to cyclic pressure variations, hydrothermal sealing, permeability reduction, and fault-valve behavior proposed by *Sibson* [1992] for earthquake rupture nucleation and recurrence, particularly in compressional tectonic regimes on faults that are otherwise unfavorably oriented for reactivation in the prevailing stress field. In either case, the crustal response to existing stress and/or strain regimes is thereby affected by production, and earthquakes may be preferentially localized beneath the producing region.

Previously, the major difficulty in recognizing the possible correlation between these events and oil production is that, in each case, the earthquakes occur at depths greater than 10 km, and considerable delays are observed between peak production and the earthquake occurrence. Peak production at Coalinga occurred in 1912; at Kettleman Hills, peak production occurred in the years 1936, 1941, 1948, 1957 and 1964 for separate sub-fields; and at Montebello, peak production occurred in 1939. It is thus hard to imagine how changes in stress related to production could have significant effects at such large distances and long time intervals. However, the apparent triggering of earthquakes at depths greater than 15 km beneath Lake Nasser 19 years after impoundment began behind the Aswan High Dam, Egypt [*Simpson et al.,* 1982], strongly suggests that such long-range interactions between the upper and lower crust do occur.

If such earthquakes are indeed considered to be triggered events, then several other earthquakes within the conterminous United States may also represent potentially induced sequences. These include: historical earthquakes near Attica, New York, in 1929, 1966 and 1967, that may be related to local solution salt mining; the extensive earthquake activity near El Reno, Oklahoma (including the 1952 $M_L$ 5.5 event), that may be related to local oil and gas production; and the 1987 $M_L$ 4.9 earthquake in southeastern Illinois that occurred adjacent to Lawrence County, one of the most productive oil regions of the State. Even the occurrence of the major damaging 1933 $M_L$ 6.3 Long Beach, California, earthquake may have been related to production in the adjacent Wilmington and Huntington Beach oil fields. The important point is that, in each of these cases, the earthquakes represent some of the largest events ever to have occurred in any of these given areas, and so represent a significant local seismic hazard. However, unlike many of the proposed, deep potentially-induced earthquakes in California [*McGarr,* 1991], these earthquakes typically exhibit focal mechanisms with a larger components of strike-slip motion, suggesting that faulting in these areas tends to occur as oblique-slip on faults at depth with steeper dips and so the spatial correlation with specific active folds may not be apparent.

## DISCUSSION AND CONCLUSIONS

Based on the case histories examined so far, it is evident that, owing to existing stress conditions in the upper crust, certain areas tend to have a relatively high probability of exhibiting seismicity related to fluid injection. In these areas, it appears that elevating formation fluid pressures by only a few megapascals (tens of bars) can trigger increased shallow seismicity. One such area prone to injection-induced seismicity is the Great Lakes region of the Appalachian Plateau, where fluid injection operations (at pressures ranging from 60 to 100 bars) have already triggered earthquakes in northeastern Ohio, western New York and southwestern Ontario (Figures 1 and 2). In most cases, the fluid-injection and the triggered earthquakes are below a regional salt layer (the Silurian Salina Formation) that acts to mechanically decouple the shallow crust (where earthquake triggering is apparently less likely) from a more critical stress state at depth [*Evans,* 1988]. Other cases of injection-induced seismicity that have occurred in Colorado (Rangely), Texas, Nebraska, Oklahoma, Arkansas, and possibly Mississippi, Louisiana and Alberta, typically involve higher injection pressures, larger volumes of injected fluid, or multiple injection wells.

The recognition that both fluid injection and the massive extraction of oil and gas may potentially trigger adjacent large earthquakes makes evaluating the seismic hazard from deep well operations much more complicated. As each of these different types of triggered events can occur up to several years after well activities have begun (or even several years after all well activities have

stopped), this suggests that the actual triggering process may be a very complex combination of effects, particularly if both fluid extraction and injection have taken place locally. In particular, it may be that, in such cases, it is the fact that *both* extraction and injection have taken place that is the major critical observation: extraction to localize strain, and injection to enhance inhomogeneities in pore fluid pressure and reduce the effective friction strength of preexisting faults [*Davis and Pennington,* 1989]. However, the important point is that models for triggered seismicity which involve local perturbations of the fluid pressure regime, or stress and strain localization resulting from massive fluid or gas withdrawal, may thus help to explain what is otherwise a rather perplexing and unusual inhomogeneous distribution of earthquakes in the central and eastern United States—an intraplate region that exhibits all the properties of fairly homogeneous tectonic environment.

Because many events induced by fluid injection are relatively shallow and may occur in close proximity to the wellbore, they pose an additional risk to both the well site and the local surrounding community. In the case of hazardous-waste disposal, there is also the potential seismic risk to the integrity of the confining layer, which if ruptured may permit the upward migration of hazardous fluid and the local contamination of potable water supplies. Because of this added risk, criteria have been established to assist in regulating well operations so as to minimize the potential for earthquake triggering by deep-well fluid injection [*Nicholson and Wesson,* 1990]. Important considerations include analyses of the hydrologic properties of the reservoir, the existing state of stress, and proximity to known or inferred fault structures. This information can then be used to help set guidelines for estimating maximum allowable injection pressures for waste disposal.

*Acknowledgements.* A major portion of this work was completed while the first author held a National Research Council Associateship with the U.S. Geological Survey in Reston, VA. This work was originally funded by the Office of Drinking Water, U.S. Environmental Protection Agency. Discussions with John Bredehoeft, Evelyn Roeloffs, Keith Evans, David Simpson, Scott Davis, and Art McGarr are much appreciated. For information regarding earthquake activity in Ohio, we would especially like to thank individuals at the Lamont-Doherty Geological Observatory, the Ohio EPA, the Ohio Geological Survey, Weston Geophysical, and Woodward-Clyde Consultants. The first author would also like to thank Steve Hickman, Rick Sibson and Ron Bruhn for the opprotunity to participate in the excellent workshop held at Fish Camp.

## REFERENCES

Allen, D.R. and M.N. Mayuga, 1970, The mechanics of compaction and rebound, Wilmington oil field, Long Beach, California, U.S.A. in *Land Subsidence,* International Association of the Science of Hydrology, *UNESCO Publ. 89,* **2,** 410–423.

Armbruster, J.G., L. Seeber and K. Evans, 1987, The July 1987 Ashtabula earthquake ($m_b$ 3.6) sequence in northeastern Ohio and a deep fluid injection well (abstract), *Seismol. Res. Lett.,* **58,** p. 91.

Caloi, P., M. DePanfilis, D. DeFilippo, L. Marcelli and M.C. Spadea, 1956, Terrimoti della Val Padana del 15–16 Maggio 1951, *Ann. Geophys.,* **9,** 63–105.

Cox, R.T., 1991, Possible triggering of earthquakes by underground waste disposal in the El Dorado, Arkansas area, *Seismol. Res. Lett.,* **62,** 113–122.

Davis, S.D., 1985, Investigations of natural and induced seismicity in the Texas panhandle, *M.S. Thesis,* University of Texas, Austin, Texas, 230 p.

Davis, S.D. and W.D. Pennington, 1989, Induced seismic deformation in the Cogdell oil field of West Texas, *Bull. Seismol. Soc. Am.,* **79,** 1477–1495.

Doser, D.I., M.R. Baker and D.B. Mason, 1991, Seismicity in the War-wink gas field, Delaware Basin, West Texas, and its relationship to petroleum, *Bull. Seismol. Soc. Am.,* **81,** 971–986.

Evans, D.M., 1966, The Denver area earthquakes and the Rocky Mountain Arsenal disposal well, *Mountain Geologist,* **3,** 23–36.

Evans, K.F., 1988, Assessing regional potential for induced seismicity from crustal stress measurements: An example from northern Ohio, *NCEER Technical Rep.,* 30 p.
Fletcher, J.B. and L.R. Sykes, 1977, Earthquakes related to hydraulic mining and natural seismic activity in western New York State, *J. Geophys. Res.,* **82,** 3767–3780.
Hamilton, D.H. and R.L. Meehan, 1971, Ground rupture in the Baldwin Hills, *Science,* **172,** 333–344.
Harding, S.T., 1981, Induced seismicity in the Cogdell Canyon Reef Oil Field, *U.S. Geol. Surv. Open-file Report 81-167,* 452–455.
Healy, J.H., W.W. Rubey, D.T. Griggs and C.B. Raleigh, 1968, The Denver earthquakes, *Science,* **161,** 1301–1310.
Hsieh, P.A. and J.D. Bredehoeft, 1981, A reservoir analysis of the Denver earthquakes: A case of induced seismicity, *J. Geophys. Res.,* **86,** 903–920.
Kovach, R.L., 1974, Source mechanisms for Wilmington Oil Field, California, subsidence earthquakes, *Bull. Seismol. Soc. Am.,* **64,** 699–711.
McGarr, A., 1991, On a possible connection between three major earthquakes in California and oil production, *Bull. Seismol. Soc. Am.,* **81,** 948–970.
Mereu, R.F., J. Brunet, K. Morrissey, B. Price and A. Yapp, 1986, A study of the microearthquakes of the Gobles oilfield area of southwestern Ontario, *Bull. Seismol. Soc. Am.,* **76,** 1215–1223.
Milne, W.G. and M.J. Berry, 1976, Induced seismicity in Canada, *Eng. Geol.,* **10,** 219–226.
Nicholson, C., E. Roeloffs and R.L. Wesson, 1988, The northeastern Ohio earthquake of January 31, 1986—Was it induced?, *Bull. Seismol. Soc. Am.,* **78,** 188–217.
Nicholson, C. and R.L. Wesson, 1990, Earthquake hazard associated with deep well injection, *U.S. Geol. Surv. Bull. 1951,* 74 p.
Pearson, C., 1981, The relationship between microseismicity and high pore pressures during hydraulic stimulation experiments in low permeability granitic rocks, *J. Geophys. Res.,* **86,** 7855–7864.
Pennington, W.D., S.D. Davis, S.M. Carlson, J. DuPree and T.E. Ewing, 1986, The evolution of seismic barriers and asperities caused by the depressuring of fault planes in oil and gas fields of south Texas, *Bull. Seismol. Soc. Am.,* **76,** 939–948.
Pratt, W.E. and D.W. Johnson, 1926, Local subsidence of the Goose Creek oil field, *J. Geol.,* **34,** 577–590.
Raleigh, C.B., J.H. Healy and J.D. Bredehoeft, 1972, Faulting and crustal stresses at Rangely, Colorado, in *Flow and Fracture of Rocks, AGU Geophys. Monogr. Series,* **16,** 275–284.
Raleigh, C.B., J.H. Healy and J.D. Bredehoeft, 1976, An experiment in earthquake control at Rangely, Colorado, *Science,* **191,** 1230–1237.
Richter, C.F., 1958, *Elementary Seismology,* W.H. Freeman and Co., San Francisco and London, 768 p.
Rothe, G.H. and C.-Y. Lui, 1983, Possibility of induced seismicity in the vicinity of the Sleepy Hollow Oil Field, southwestern Nebraska, *Bull. Seismol. Soc. Am.,* **73,** 1357–1367.
Segall, P., 1989, Earthquakes triggered by fluid extraction, *Geology,* **17,** 942–946.
Sibson, R.H., 1992, Implications of fault-valve behavior for rupture nucleation and recurrence, *Tectonophysics,* **211,** 283–293.
Simpson, D.W., 1986, Triggered earthquakes, *An. Rev. Earth Planet. Sci.,* **14,** 21–42.
Simpson, D.W., R.M. Kebeasy, M. Maamoun, R. Albert, and F.K. Boulois, 1982, Induced seismicity at Aswan Lake, Egypt (abstract), *Eos, Trans. Am. Geophys. Union,* **63,** p. 371.
Simpson, D.W. and W. Leith, 1985, The 1976 and 1984 Gazli, U.S.S.R, earthquakes—Were they induced?, *Bull. Seismol. Soc. Am.,* **75,** 1465–1468.
Teng, T.L., C.R. Real and T.L. Henyey, 1973, Microearthquakes and water flooding in Los Angeles, *Bull. Seismol. Soc. Am.,* **63,** 859–875.
Wetmiller, R.J., 1986, Earthquakes near Rocky Mountain House, Alberta, and their relationship to gas production facilities, *Canadian J. Earth Science,* **23,** 172–181.
Yerkes, R.F. and R.D. Castle, 1976, Seismicity and faulting attributed to fluid extraction, *Eng. Geol.,* **10,** 151–167.

# THE ROLE OF CRUSTAL FLUIDS IN THE TRIGGERED RESPONSE OF LONG VALLEY CALDERA TO THE M=7.3 LANDERS, CALIFORNIA, EARTHQUAKE

by

David P. Hill, Malcolm J.S. Johnston, John O. Langbein, and *Jeffrey Behr

U.S. Geological Survey
Menlo Park, CA 94025

*CIRES, University of Colorado
Boulder, CO 80302

The M = 7.3 Landers, California, earthquake of 28 June 1992 triggered an abrupt increase in seismicity at widely scattered sites across much of the western United States (Hill and others, 1993). Most of the sites of triggered seismicity were located north-northwest of the Landers epicenter (the direction of rupture propagation) along the eastern California shear zone (the Sierra Nevada-Great Basin boundary zone SNGBZ) and its extension into the southern-most Cascade volcanoes, Lassen Peak and Mount Shasta. Outlying sites of triggered activity include the Geysers geothermal area in the Coast Ranges of northern California, an area near Cedar City Utah along the Hurricane fault zone, an area along the western margin of the Idaho batholith near Cascade, Idaho, and a site in Yellowstone National Park (this was the most distant site of recognized triggered seismicity at 1250 km or 17 source dimensions from the Landers epicenter). The remotely triggered seismicity was confined to areas characterized by persistent seismicity and strike-slip to normal faulting (a horizontal least principal stress). Many of the remotely triggered sites are areas of geothermal and recent volcanic activity. Linear elasticity theory applied to laterally homogeneous earth models fails to provide an explanation for the triggering of such remote seismicity. We must thus look to possible non-linear interactions of the static and dynamic elastic fields with local crustal heterogeneities to understand the triggering process. Some of the most promising models involve interactions of the large dynamic strains radiated from the mainshock with fluids (including possibly magma) in volumes of the crust producing triggered seismicity.

Of the fifteen recognized sites with seismicity remotely triggered by the Landers mainshock, only Long Valley caldera is covered by both a dense seismic network and continuous deformation-monitoring instrumentation. We use this unique data to study response of Long Valley caldera to the Landers earthquake for clues to the processes driving the remotely triggered activity. We focus in particular on the spatial-temporal evolution of the triggered seismicity and its relation to a transient strain pulse detected by the deformation monitoring instrumentation in the caldera (Figure 1).

Long Valley caldera is just over 410 km north-northwest of the Landers mainshock epicenter, and it is roughly midway along the eastern California shear zone. The caldera was formed by the massive eruption of the Bishop tuff some 740,000

years ago. It has since been the site of recurring volcanic eruptions (the most recent 500 ago) and, through historic time, persistant earthquake activity (Bailey and Hill, 1990). Current unrest within the caldera includes recurring earthquake swarm activity accompanied by inflation of the resurgent dome with a cumulative uplift of over 60 cm since its onset in 1979-1980 (Langbein and others, 1993).

Figure 2 summarizes the spatial-temporal evolution of the seismicity within and immediately south of the caldera triggered by the Landers mainshock (origin time: 1158 UT on 28 June 1992). The triggered seismicity occupies essentially the same crustal volume as the recurring swarm activity that began in 1980. The first earthquakes triggered within the caldera followed the Landers S wave by approximately 40 seconds (about 3 minutes after the Landers origin time [AL]) and occurred during passage of the crustal Love and Rayleigh waves, which carried the peak dynamic stresses (approximately 3 bars in the vicinity of Long Valley; Michael, 1992; Hill and others, 1993). This initial triggered activity was located at a depth of 6 to 8 km beneath the southwestern margin of the resurgent dome in the same volume that has produced the most frequent and intense swarm activity since 1980. This initial activity included a M=2.2 event 16 minutes AL and a M=2.6 event 32 minutes AL at depths of 7 and 8 km respectively. By 34 minutes AL, triggered seismicity began near the south east margin of the resurgent dome some 6 km east of the initial activity. Effectively, much of the southern part of the caldera became activated within half an hour of the Landers mainshock with no evidence for a migration front from the site of initial activity. This stands in contrast to typical swarm activity in the caldera, which tends to be concentrated in isolated clusters within the seismogenic zone with only one cluster active at a time. Only the largest swarms (M > 5 events) involve activity throughout the seismogenic southern half of the caldera.

Over the next 17 hours, seismicity gradually became concentrated in a north-trending band south of the eastern margin of the resurgent dome (between longitude 118 50 and 51). A M=3.7 earthquake along the southern extension of this zone 3 km south of the caldera that occurred at 0537 UT the next day (June 29). This was the largest event in the vicinity of the caldera during the triggered activity, and it appears to coincide with renewed activity throughout the southern part of the caldera (secondary triggering?). Activity throughout the area gradually returned to background seismicity rates over the next week (Figure 2). A preliminary look at focal mechanisms for the triggered seismicity indicates no significant differences from the long term swarm activity in the caldera, which is dominated by strike-slip to normal-oblique mechanisms and T-axes with a northeast-southwest orientation. We have more work to do on the focal mechanism issue, however.

The largest and best located of the triggered earthquakes are located near the base of the seismogenic crust, which varies in depth from 12 to 15 km in the Sierra Nevada block south of the caldera to less than 6 km beneath the resurgent dome (Hill, 1992). The depth-time plot in Figure 2 admits the possibility that the M > 1.5 triggered seismicity migrated to shallower depths at a rate of roughly 6 km/day from its initiation at depths of 6 to 8 km beneath the south moat of the caldera.

The borehole dilatometer POPA located 3 km west of the caldera (Figure 2) recorded a compressional strain transient that began during the S wave arrival from the Landers mainshock, grew to 0.25 microstrain over the next five days, and decayed to

near background level over the next two to three weeks (Figure 1). A 0.2 microradian down-to-the-east tilt transient with a similar time history was recorded by the long base tiltmeter located near the southeastern margin of the resurgent dome. Repeated measurements of two-color geodimeter network spanning the caldera during this time interval detected no strain response to the Landers earthquake above the 0.2 microstrain resolution of the measurements. These measurements indicate that the Landers earthquake triggered a broadly distributed pressure increase beneath the caldera that gradually decayed to background over the next several weeks. Although the available deformation data are too sparse to resolve the areal extent of the pressure source, it is clear that they can not be satisfied solely by pressurization of the concentrated (Mogi) source 7 to 10 km beneath the center of the resurgent dome that satisfies most of the resurgent dome deformation data since the onset of deformation in 1979-1980.

The close correlation between the plot of cumulative seismicity with time and the compressional phase of the strain transient following the Landers mainshock (Figure 2) suggests a causal link that most likely involves a transient increase in fluid pressures within the seismogenic crust beneath the caldera that gradually decay by diffusion over a period of a few weeks. Both observed and theoretical static strain changes from the Landers mainshock seem too small to produce such a response (the static strain step recorded by the POPA dilatometer was 0.003 microstrain consistent with the static strain field predicted for the Landers dislocation in an elastic half space). The large dynamic strains focused toward Long Valley caldera (12 microstrain recorded in the vicinity of the caldera) and the eastern California shear zone by strong north-northwestward directivity of the Landers source admit several plausible mechanisms for triggering such a transient pressure increase. Two of the more interesting possibilities that may apply to Long Valley caldera include:

1) Pumping mid-crustal pore fluids into the overlying seismogenic crust by the dilatational component of the crustal Rayleigh waves. Water levels in wells tapping unconfined aquifers often fluctuate greatly during the passage of Rayleigh waves from large, distant earthquakes. Pore pressure oscillations driven by Rayleigh waves acting on confined crustal fluids may rupture fluid seals releasing high-pressure fluids into adjacent crustal volumes by processes similar to those described by Byerlee (1993) for fault zones. For example, near-lithostatic pore fluids sealed in the brittle-plastic transition zone may be driven into the overlying seismogenic crust by this process.

2) Accelerating the exsolution of volatile components in subadjacent magma bodies thereby increasing the pressure within the magma bodies (Sahagian and Proussevitch, 1992) or increasing pore pressure in the overlying rock by an increased flux of volatiles from the magma bodies.

The observation that the triggered seismcity in Long Valley caldera began near the base of the seismogenic crust and that most of the moment release was concentrated there is consistent with both possibilities.

## REFERENCES

Byerlee, J., 1993, A model for episodic flow of high pressure water in fault zones before earthquakes; Geology (in press).

Bailey, R.A, and D.P. Hill, 1990, Magmatic unrest at Long Valley caldera, California, 1980-1990, Geoscience Canada, v. 17, pp. 175-179.

Hill, D.P., 1992, Temperatures at the base of the seismogenic crust beneath Long Valley caldera, California, and the Phlegrean Fields caldera, Italy; in P. Gasparini, R. Scarpa, and K. Aki (eds) Volcanic Seismology, Springer-Verlag, New York, pp. 432-461.

Hill, D.P., P.A. Reasenberg, A. Michael, and 28 others, 1993, Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, Science (in press).

Langbein, J.O., D.P. Hill, T.N. Parker, and S.K. Wilkinson, 1993, An episode of reinflation of the Long Valley caldera, eastern California; 1989-1992; Jour. Geophys. Res. (in press)

Michael, A., 1992, Initiation of seismicity remotely triggered by the Landers earthquake: where and when (abs); EOS, v. 73, p. 392-393.

Sahagian, D.L. and A.A. Proussevitch, 1992, Advective bubble overpressure in volcanic systems (abs); EOS, v. 73, p. 627-628.

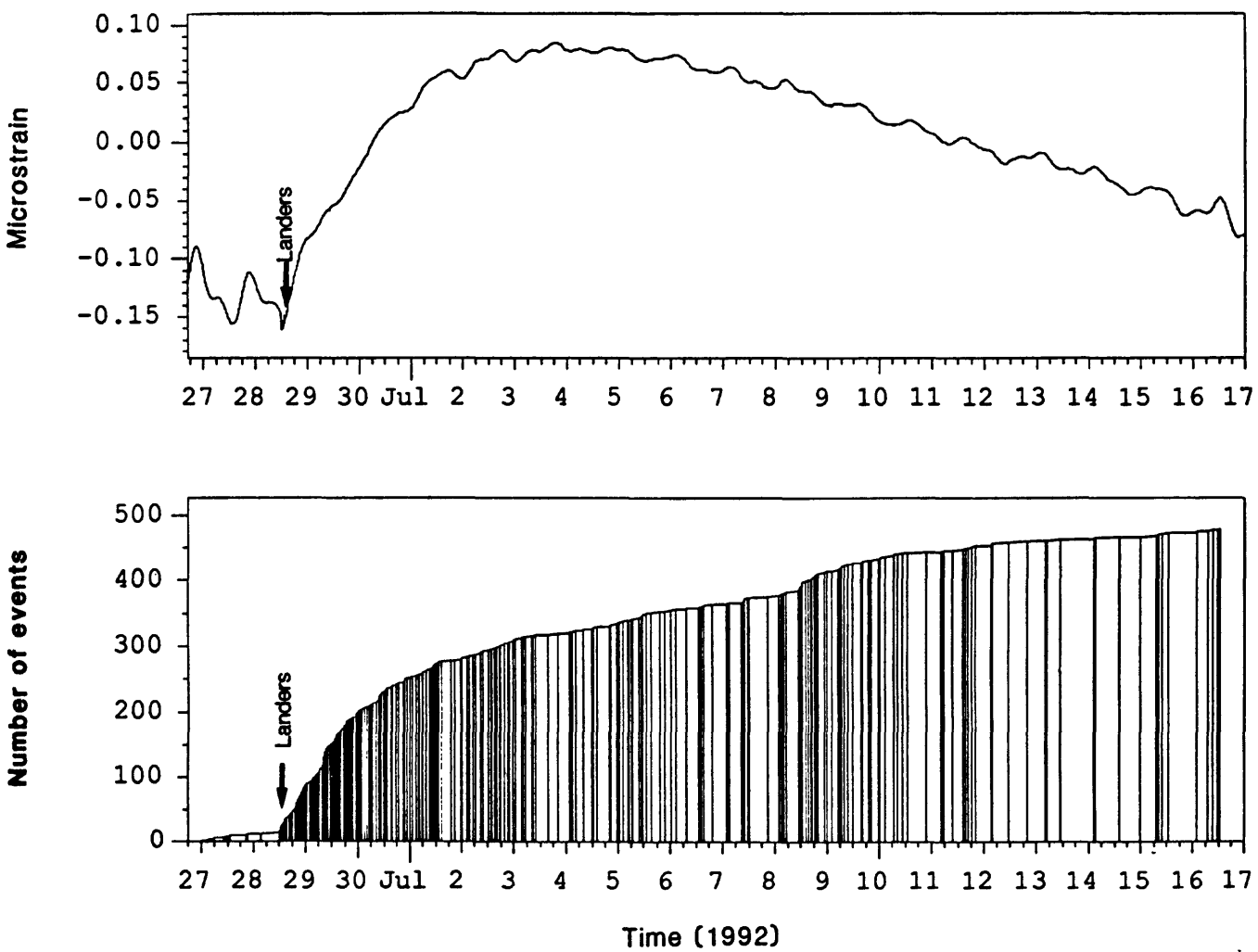



Figure 1. TOP: Postpile (POPA) borehole dilatometer strain from 27 June through 17 July, 1992. BOTTOM: Cumulative number of earthquakes with time from 27 June through 17 July, 1992. Each vertical line represents an earthquake.

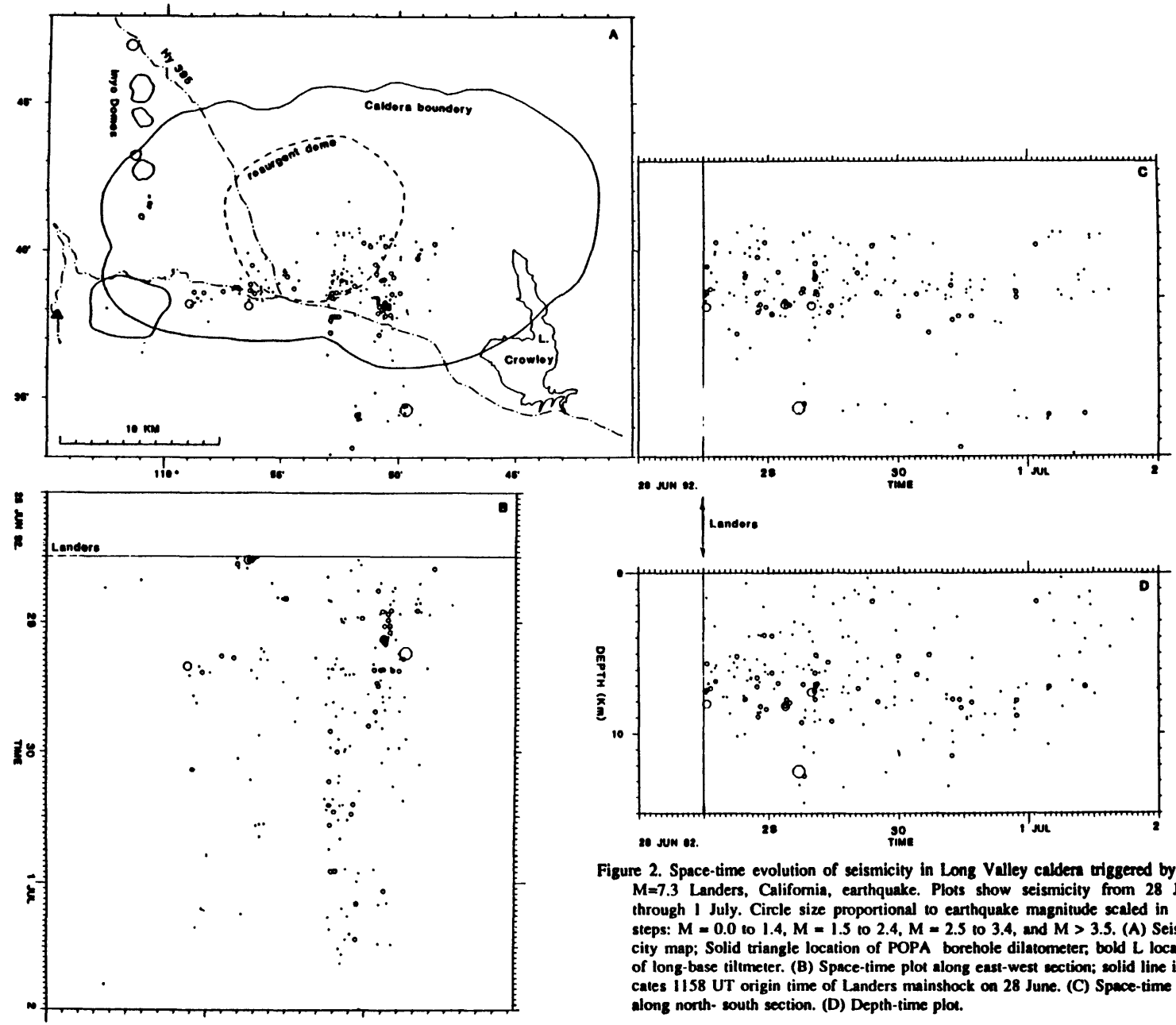



Figure 2. Space-time evolution of seismicity in Long Valley caldera triggered by the M=7.3 Landers, California, earthquake. Plots show seismicity from 28 June through 1 July. Circle size proportional to earthquake magnitude scaled in four steps: M = 0.0 to 1.4, M = 1.5 to 2.4, M = 2.5 to 3.4, and M > 3.5. (A) Seismicity map; Solid triangle location of POPA borehole dilatometer; bold L location of long-base tiltmeter. (B) Space-time plot along east-west section; solid line indicates 1158 UT origin time of Landers mainshock on 28 June. (C) Space-time plot along north- south section. (D) Depth-time plot.

# REVERSED-POLARITY SEISMIC REFLECTIONS ALONG FAULTS OF THE OREGON ACCRETIONARY PRISM: INDICATORS OF FLUID MIGRATION AND ACCUMULATION

**By**

**J. C. Moore, G. R. Cochrane, Earth Sciences, UC Santa Cruz**
**Santa Cruz, CA, 95064**
**G. F. Moore, SOEST, University of Hawaii,**
**Honolulu, Hawaii, 96822**

## Abstract

High amplitude reversed-polarity reflections occur commonly on the frontal thrust and in the proto-thrust zone of the Oregon accretionary prism. The highest amplitude reversed-polarity reflections occur on the frontal thrust. The reversed-polarity reflections on the proto-thrusts are distinct but of lower amplitude than those from the frontal thrust.

The reversed-polarity reflections could be caused by either dilation of the fault zone or thrusting of higher over lower impedance sediment. Reversed-polarity reflections on the proto-thrusts are probably not due to thrust inversion because the vertical displacement of proto-thrusts averages only 27 m and the occurrence of polarity reversals is uncorrelated to displacement magnitude. The predicted reflection coefficient due to the 27 m mean offset of a velocity-derived impedance curve is an order of magnitude less that the observed reflection coefficient of the proto-thrusts. Because the polarity reversals are randomly distributed in depth throughout the proto-thrusted section it is unlikely that they arise from a thrust inversion of any stratigraphically controlled variation in density and velocity. The reversed-polarity reflections from the proto-thrusts are probably due to fault zone dilation.

The 1.5 km throw on the frontal thrust suggests a thrust inversion could produce the observed reflections. Because the high amplitude reversed-polarity reflections lie in bands at relatively shallow depths, they could be produced by juxtaposition of high impedance sediments with stratigraphically controlled bands of low impedance sediments. Thus, some reversed-polarity reflections are indicative of dilation and others might be produced by thrusting.

Fluid dilating the proto-thrusts may be derived from a rapidly deposited turbidite fan sequence or a subjacent overpressured abyssal plain turbidite sequence. The detachment or "proto-decollement" at the base of the proto-thrust zone occurs in an interval of anomalously low velocity, high porosity, and high fluid pressure; this interval may be a fluid source zone. At the base of the proto-thrust zone, clay mineral dehydration and hydrocarbon generation are nearing completion, which may also contribute to overpressuring.

We believe that fluids are sourced in the overpressured sequences and migrate up the steep proto-thrusts with a relatively rapid drop in overburden pressure with respect to head loss. The excess fluid pressure ultimately drives the fault zone into tensile failure. If the compressional state of stress induces failure prior to dilation, the decreased differential stress would allow tensile failure if the rate of fluid flow and local pore pressure increase exceed the stress buildup.

## Introduction

Seismic reflection images of faults from modern accretionary prisms are commonly marked by high amplitude reversed-polarity reflections (Shipley et al., 1990)( Bangs and Westbrook, 1991)( Moore and Shipley, 1993)( Cochrane et al., 1994). The reversed-polarity reflections indicate a decrease in acoustic impedance (the product of velocity and density) across the faults. All of these reversed-polarity reflections occur on decollement zones that thrust highly deformed deposits of the accretionary prisms over undeformed sediments. Thus, thrusting higher impedance over lower impedance material provides a simple method to develop a reversed-polarity reflection. Alternatively, dilation of the fault zone could reduce its velocity and density relative to the wall rocks, create a decrease in impedance, and generate a reversed-polarity reflection.

Because dilated fault zones have substantially increased permeability, and channelize fluid flow, it is important to distinguish if the reversed-polarity reflections indicate dilation or whether they are simply the results of thrusting. If the reversed-polarity reflections are dilated zones of low effective stress then they could substantially reduce the fault strength; moreover, they may represent analogs of the fluid compartments hypothesized along the San Andreas Fault (Byerlee, 1993).

Wave-form models supporting both mechanisms for producing high amplitude reversed-polarity reflections have been published. A model of lower impedance sediment underthrusting higher impedance deformed accretionary prism successfully reproduces the reversed-polarity reflection of the decollement zone off Costa Rica (Shipley et al., 1990). Along the decollement zone of the Northern Barbados Ridge, Bangs and Westbrook (1991) best-fit the reversed-polarity reflection with a zone 20 m thick with a 10% velocity reduction. The reversed-polarity reflection of the Nankai Trough decollement zone was modeled either as a zone of reduced velocity or as a simple impedance inversion (Moore and Shipley, 1993).

Here we evaluate high amplitude reversed-polarity reflections from the Oregon margin from a structural and geometric perspective. We utilize multichannel seismic reflection data collected in 1989 by Digicon and processed at the University of Hawaii (MacKay et al, 1992; Cochrane et al., 1993). Overall this survey collected 2000 line-km of data in a 4700 $km^2$ area; our study concentrates on a limited grid in a 70 km2 area (Fig. 1).

Fault geometry was been defined using seismic depth sections. The depth conversion utilized the stacking velocities from seismic processing, in combination with refraction data and were adjusted to maintain a uniform slope to the oceanic crust as it projects beneath the first ridge (Cochrane et al., 1993). The depth sections have been examined with seismic interpretation and visualization software from Landmark Graphics Corporation.

## Tectonic Setting of the Oregon Margin

The Oregon-Washington subduction zone is characterized by copious clastic sediment influx and a young oceanic plate. A topographic trench is absent nearly everywhere as two large submarine fans dominate the depositional system. Our research is focused off central Oregon on the southern fringe of the Astoria Fan (Fig. 1). Here, the oceanic crust is about 9 Ma old, converges at a rate of approximately 4 cm/year along 070°, and is covered by nearly 4 km of sediment (Kulm et al., 1984) (DeMets et al., 1990). This area of central Oregon contains distinct domains of landward and seaward verging structures (Kulm et al., 1973) (Snavely and Miller, 1986) (MacKay et al., 1992). Our

study concentrates in a province of seaward vergent structures; a transition to landward vergent thrust faults occurs approximately 20 km to the north of the area we worked.

## Fault Geometry of the Frontal Deformation Zone

The central portion of the seaward vergent province (Figs. 1,2) is characterized by a distinct first ridge formed by a fault-bend fold lying above the frontal thrust. The frontal thrust elevates the first ridge about 800 m above the adjacent abyssal plain and has a throw of about 1.5 km. The frontal thrust dips at 20 degrees to about 4 km below sea level and then steepens to about 40 degrees until it intersects the decollement at about 5.5 km below sea level (Fig. 3). The geometry of the frontal thrust is consistent for at least the 4 km along strike transected by lines 4-9 (Fig. 3). Fan deposits dredged from the first ridge contain radiolarians < 0.3 Ma old (Kulm and Fowler, 1974), providing an upper limit for the age of the frontal thrust.

Extending about 6 km seaward of the frontal thrust, the proto-thrust zone encompasses numerous small-scale faults that cut the incoming section. Proto-thrusts dip uniformly at about 40 degrees (Fig. 3). Individual proto-thrusts cannot be followed north-south from line to line (Fig. 3); they apparently are limited in lateral extent in addition to throw. Throw has been determined by mapping a horizon across seismic lines 5-9 and measuring the vertical offset at each proto-thrust (Fig. 4). The mean throw of the proto-thrusts is 27 m with extraordinary values to 100 m (Fig. 5).

## Nature and Distribution of High Amplitude Reflections from Faults

Reflection polarity was analyzed by displaying our seismic data in color on a workstation using interpretation software from Landmark Graphics Corporation. We worked with relative amplitude data that had been corrected only for spherical divergence. The normal-polarity, water-bottom reflection is a positive-negative-positive composite wavelet whereas the reversed-polarity reflection is the opposite. Only high amplitude reversed-polarity reflections were only mapped in this study.

Seismic lines across the frontal thrusts commonly show discontinuous high amplitude reversed polarity reflections. Two bands of polarity reversals extend from near the surface to more than halfway down the frontal thrust. Individual reversed-polarity reflections have a mean down-dip extent of 1 km (Fig. 6). An along strike projection (Fig. 7) shows that these high amplitude reflections correlate well, though the perspective view (Fig. 8) illustrates changes in down-dip extent along strike.

Polarity reversals on the proto-thrusts are lower amplitude than those on the frontal thrusts, are shorter in down-dip extent, and do not correlate from line to line. The seismic lines are separated by about 0.8 km on average and the down-dip extent of the reversals has a mean value of 0.5 km. Therefore, the extent of the reversal on the proto-thrust fault surface along strike cannot be more than about 3 times its down-dip extent.

A north-south projection of all reversals on to an east-west line shows that the reversals don't preferentially occur in any depth interval or series of depth intervals (Fig. 7). Apparently, they are not stratigraphically controlled; for example, a particular unit of low acoustic impedance is not overthrust by a subjacent interval of higher acoustic impedance, therefore generating a band of reversed polarity reflections at a uniform depth. The occurrence of polarity reversals shows no correlation with proto-thrust offset (Fig. 5) and therefore is unlikely to be a simple result of impedance inversion due to fault displacement.

In summary, the distributed occurrence of the polarity reversals on proto-thrusts of very small vertical displacement and their lack of correlation to fault offset suggests that the reversals are dilational in origin. Geological arguments cannot be made for dialational origin of the polarity reversals on the frontal thrust, because they lie along two bands and are on a fault with substantial throw. However, our results don't exclude dilational origin for these reflections

## Geophysical Arguments for the Origin of the Polarity Reversals

Wave-form modeling of high amplitude reversed-polarity reflections from decollement or proto-decollement zones have been successful at a number of localities (Shipley et al., 1990; Bangs and Westbrook, 1991; Moore and Shipley, 1993; Cochrane et al., 1994). In all of these examples the reflections dip a few degrees and the reflections can be modeled as horizontal surfaces. Off Oregon, the proto-thrusts and frontal thrust dip 20 to 40 degrees making their waveform modeling problematic.

To investigate the possible origin of the high amplitude reflections from proto-thrusts, we estimated reflection coefficients of the seafloor and the proto-thrusts (White, 1977). This approach relates the reflection coefficient to the amplitude of the seafloor reflection and the amplitude of the first seafloor multiple. Using 75 traces from an unmigrated, near-trace section, we found the mean value of seafloor reflection coefficient to be 0.202. Making only minimal corrections for spherical divergence and transmission loss, and ignoring absorption, we calculated the reflection coefficient of the proto-thrusts to be approximately -0.04 to - 0.06. These minimum values are 20-30% of the seafloor reflection coefficient.

Offsets of the stratified section along the proto-thrust predict reflection coefficients an order of magnitude less than those observed in the reflection data. An impedance curve for the stratigraphic section was calculated from interval velocities (Cochrane et al., 1994) and densities derived therefrom. The smoothed impedance curve was then offset 27 m or the mean displacement of the proto-thrust resulting in a small impedance change; The reflection coefficient predicted by fault displacement is determined by taking the difference in impedance at a point caused by the displacement divided by the sum of the impedance above and below the interface. The predicted value is -0.003; even if the offset were increased to the maximum fault displacement of about 100 m, the impedance contrast would increase four times and still only be a fifth to a third of the observed values. Because the amplitude of the reversed-polarity reflections in the proto-thrusts cannot be generated by simple fault offset, the reflections may be of dilational origin.

## Conditions Conducive to the Production of High Fluid Pressures.

Both the geological and geophysical evidence favor production of the reversed-polarity reflections on proto-thrusts by dilation of the fault zone. Because the sediments are underconsolidated and weak, the dilation must be due to high fluid pressures and not simply volumetric expansion caused by the rotation of irregular rigid particles. The geological setting of the proto-thrust zone is conducive to the development of high fluid pressures

The velocity data derived from the seismic processing (Cochrane et al., 1994) have been converted to porosities using relationship from (Hamilton, 1978) (Fig. 9). This porosity curve decreases systematically from the surface to near the top of the abyssal plain turbidite unit where it increases down section. This change in the porosity gradient is indicative of an overpressured sequence (Fertl, 1976).

**High Sedimentation Rates:** The incoming sedimentary section is a rapidly-accumulated turbidite fan sequence in which high fluid pressures would be expected for depositional reasons alone. The incoming sedimentary sequence is 3.5 km thick over oceanic crust about 8.75 my old (Kulm et al., 1984). The sedimentation rate for the abyssal plain turbidite (Fig. 9) is estimated to be modestly higher than the maximum value measured for the equivalent unit at Site 174 (about 70 km seaward of the margin) (Kulm and von Huene,1973). Projection of this rate (250 m/my) upward from the oceanic crust to the base of the of the turbidite fan deposits suggests the onset of turbidite fan deposition at about 1.75 ma and a rate of deposition of 1000 m/my. Onset of the fan sedimentation at this time (early Pleistocene) is consistent with the belief that the fans are Pleistocene features. This more rapid pulse of fan sedimentation would both load the underlying more mud-rich turbidites and inhibit the dewatering of the fan sediments themselves. In both cases high overpressures are likely to result.

**Mineral Dehydration and Hydrocarbon Generation:** Both clay mineral dehydration (smectite to illite transition) and the maturation of organic matter produce fluids and would contribute to the overpressuring of the incoming sedimentary section off Oregon. Both of these reactions are primarily temperature dependent. The thermal gradient in Cascadia Basin is about 60 degrees per kilometer as estimated from surface heat flow measurements (A. Fisher pers. com.), and drilling (Carson, Westbrook, Musgrave, et al., 1993) (Fig. 9) The thermal gradient may be slightly higher at depth because of the blanketing effects of the rapid deposition of the turbidite fan. Using 60 degrees/km will conservatively evaluate the effects of these processes.

The curve of clay mineral dehydration in Figure 9 is derived from kinetic equations (Pytte and Reynolds 1989) and shows that substantial proportion of the dehydration would be complete at the depth of the upper part of the abyssal plain turbidite sequence; this is the depth at which the velocity-porosity data suggest overpressuring and the proto-decollement occurs. At this depth hydrocarbon maturation would near its peak (Tissot and Welte, 1984) further contributing to overpressures.

**Tectonic Deformation:** The conjugate nature of some of the proto-thrusts indicates a maximum principal stress oriented subhorizontal. This lateral shortening of the sedimentary section may contribute to overpressuring by lateral consolidation. The increase in velocity laterally as the sedimentary section is transported into the proto-thrust zone supports this hypothesis (Cochrane et al., 1994).

**Discussion:** Depositional, diagenetic, and tectonic aspects of the sediments in the proto-thrust zone all suggest they are overpressured. The subhorizontal orientation of the maximum principal stress in the proto-thrust zone suggests that the shear stress is very low on the proto-decollement and consequently the fluid pressure is very high (Seely, 1977).

## Geomechanical-Hydrological Model:

The thick incoming sedimentary sequence provides a viable source of high pressure fluids to dilate the proto-thrusts and create the reversed polarity reflections. We envision the following conceptual model for this process (Fig. 10): 1) The fluids are sourced in an overpressured interval. 2) The fluids migrate up the steep proto-thrusts with a relatively rapid drop in overburden pressure with respect to head loss. 3) The excess fluid pressure ultimately drives the fault zone into tensile failure. If the compressional state of stress induces failure prior to dilation, the decreased differential stress would allow tensile failure if the rate of fluid flow and local pore pressure increase exceeded the stress buildup. If this model is correct, then the occurrence of polarity reversals over a wide range of depth intervals suggests that fluid is not being sourced from a single zone of uniform

overpressure; rather fluid is probably migrating from a number of overpressured intervals, both in the abyssal plain turbidite sequence and in the turbidite fan sequence.

This model requires substantial fluid migration up the faults at sub-lithostatic fluid pressures, prior to reaching the hydrofracture threshold and creation of the reversed-polarity reflection. The hydrofracture condition could provide ample permeability for fluid flow; however, the relationship between permeability and effective stress for fault rocks in accretionary prisms needs to be established for pressures just sublithostatic.

## Acknowledgements

Funded by NSF grants OCE8813907 and OCE9116368 to J.C.M., and grant OCE-8821577 to G.F.M. Seismic interpretation software provided by University Parthership Program of Landmark Graphics Corporation.

## Bibliography

Bangs, N. L. B., and Westbrook, G. K., 1991, Seismic modeling of the decollement zone at the base of the Barbados Ridge accretionary complex: Journal of Geophysical Research, v. 96, p. 3853-3866.

Byerlee, J., 1993, Model for episodic flow ofhigh-pressure water in fault zones before earthquakes: Geology, v. 21, p. 303-306.

Carson, B., Westbrook, G., Musgrave, R., and others, 1993, Proceedings of the Ocean Drilling Program: Initial Results, College Station, TX (Ocean Drilling Program), v. 146, p. __p.

Cochrane, G., MacKay, M., Moore, G. F., and others, 1993, Multichannel seismic survey of the central Oregon margin: In Carson, B., Westbrook, G., Musgrave, R., et al., Proceedings Ocean Drilling Program, Initial Reports: College Station, TX (Ocean Drilling Program), v. 146, p. __.

Cochrane, G. R., Moore, J. C., MacKay, M. E., and others, 1994, Velocity-porosity model of the Oregon accretionary prism from seismic reflection and refraction data: Journal of Geophysical Research, v. __, p. __.

DeMets, C., Gordon, R. G., Argus, D. F., and others, 1990, Current plate motions: Geophys. J. Int., v. 101, p. 425-478.

Fertl, W. H., 1976, Abnormal formation pressures: New York, Elsevier Scientific Publishing Company, 382 p.

Hamilton, E. L., 1978, Sound velocity-density relations in sea-floor sediments and rocks: Journal of the Acoustical Society of America, v. 63, p. 366-377.

Kulm, L. D., and Fowler, G. A., 1974, Oregon continental margin structure and stratigraphy: A test of the imbricate thrust model: In The Geology of Continental Margins, C. A. Burke and C. L. Drake, (Eds.), New York: Springer-Verlag, p. 261-284.

Kulm, L. D., Peterson, C. P., Connard, G. G., and others, 1984, Western North American continental margin and adjacent ocean floor off Oregon and Washington: Atlas 1 Ocean Margin Drilling Program, Regional Atlas Series: Woods Hole, MA, Marine Sciences International, 32 sheets.

Kulm, L. D., Prince, R. A., Snavely, P. D., and others, 1973, Site survey of the northern Oregon continental margin and Astoria Fan, in Kulm, L. D., et al.,: Initial Reports of the Deep Sea Drilling Project, v. XVIII: Washington, D.C., U. S. Government Printing Office, p. 979-986.

Kulm, L. D., and von Huene, R. e. a., 1973, Initial Reports of the Deep Sea Drilling Project: v. 18: Washington (U.S. Government Printing Office), 1077 p.

MacKay, M. E., Moore, G. F., Cochrane, G. R., and others, 1992, Landward vergence and oblique structural trends in the Oregon margin accretionary prism: implications and effect on fluid flow: Earth and Planetary Science Letters, v. 109, p. 477-491.

Moore, G. F., and Shipley, T. H., 1993, Character of the decollement in the Leg 131 area, Nankai Trough: In Hill, I. A., Taira, A., Firth, J. V. et al., Proc. Ocean Drilling Program, Scientific Results, College Station, TX, v. 131, p. 73-82.

Pytte, A. M., and Reynolds, R. C., 1989, The thermal transformation of smectite to illite: in Naeser, N. D. and McCulloh, eds., Thermal history of sedimentary Basins: methods and case histories, Springer-Verlag (New York), p. 33-40.

Seely, D. R., 1977, The significance of landward vergence and oblique structural trends on trench inner slopes: in Talwani, M., and Pitmann, S. C., eds., Island Arcs, Deep Sea Trenches, and Back-Arc Basins, American Geophysical Union, Maurice Ewing Series 1, p. 187-198.

Shipley, T. H., Stoffa, P. L., and Dean, D. F., 1990, Underthrust sediments, fluid migration paths and mud volcanoes associated with the accretionary wedge off Costa Rica: Middle America Trench: Journal of Geophysical Research, v. 95, p. 8743-8752.

Snavely, P. D. J., and Miller, J., 1986, The central Oregon continental margin, lines WO76-4 and WO76-5; in Roland von Huene, ed., Seismic Images of Modern Convergent Margin Tectonic Structure: American Association of Petroleum Geologists Studies in Geology #26, p. 24-29.

Tissot, B. P., and Welte, D. H., 1984, Petroleum Formation and Occurrence: New York, Springer-Verlag, 699 p.

White, R. S., 1977, Seismic bright spots in the Gulf of Oman: Earth and Planetary Sciences Letters, v. 37, p. 29-37.

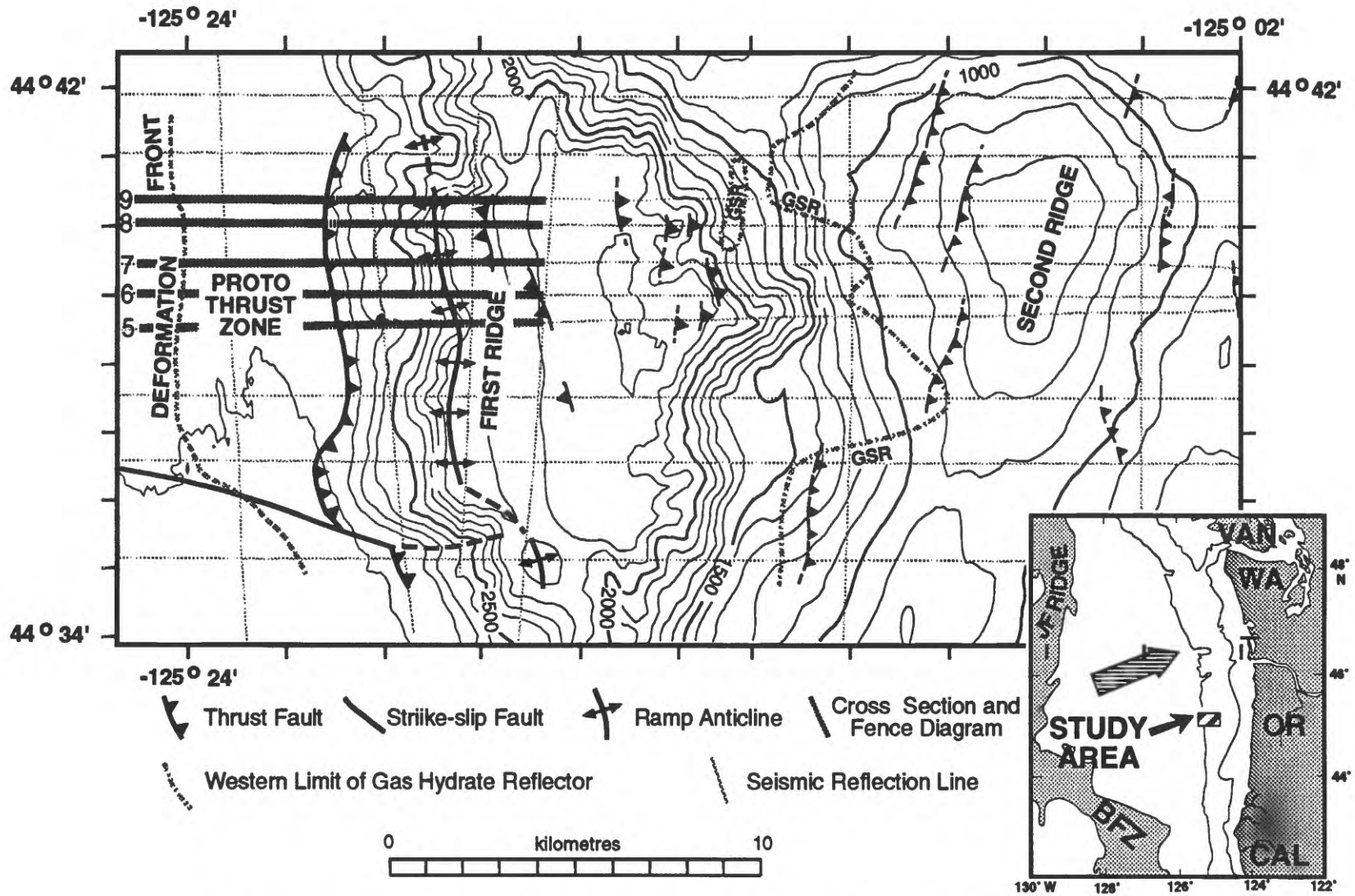



Figure 1. Bathymetric map and major structural features of seaward vergent portion of Oregon margin. Modified from After MacKay et al., 1992.

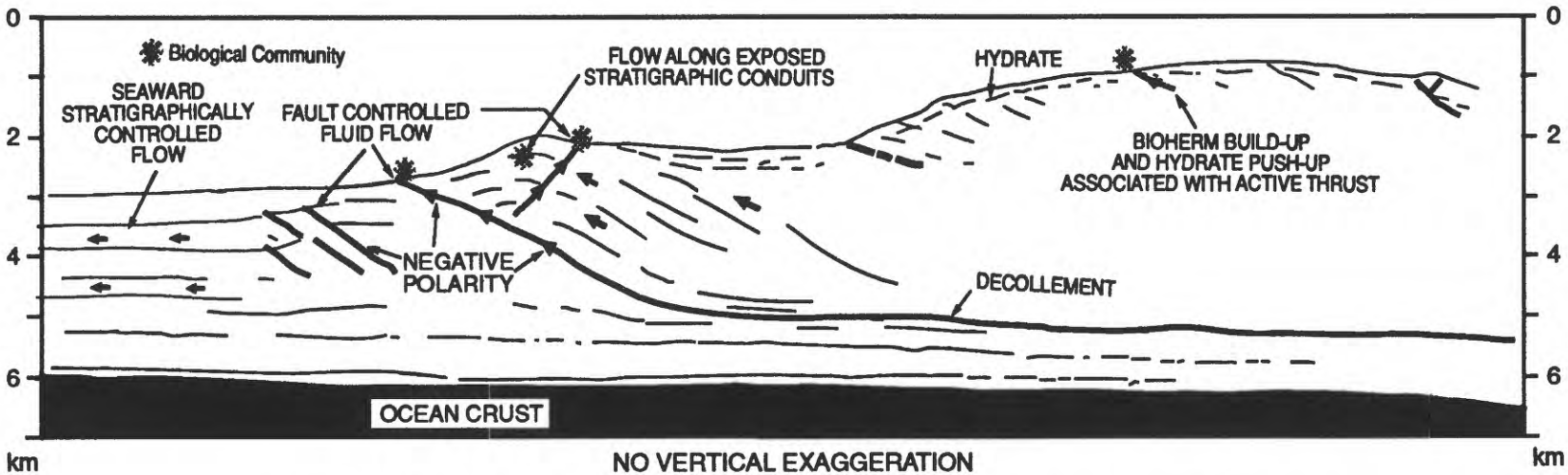



Figure 2. Generalized cross section through the seaward vergent portion of the Oregon accretionary prism.

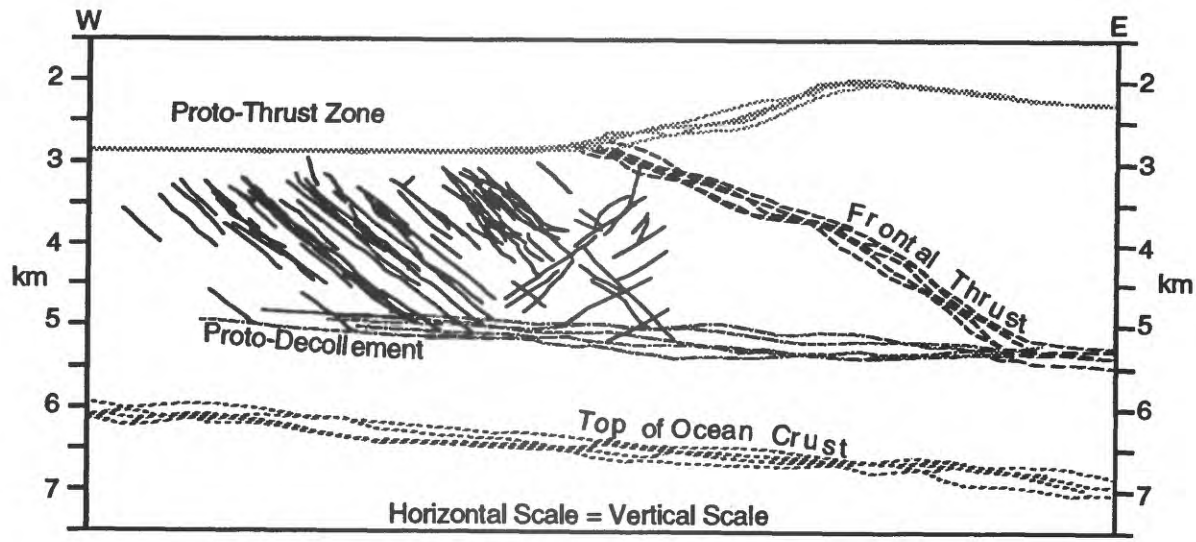



Figure 3. Geometry of principal structural features of frontal deformation zone from lines 5-9 as viewed in composite E-W section (view south to north. Note conjugate nature of proto-thrusts. Seismic lines trending north-south crossing these sections indicate that frontal thrust is continuous. Available data suggests proto-thrusts are not continuous between seismic lines.

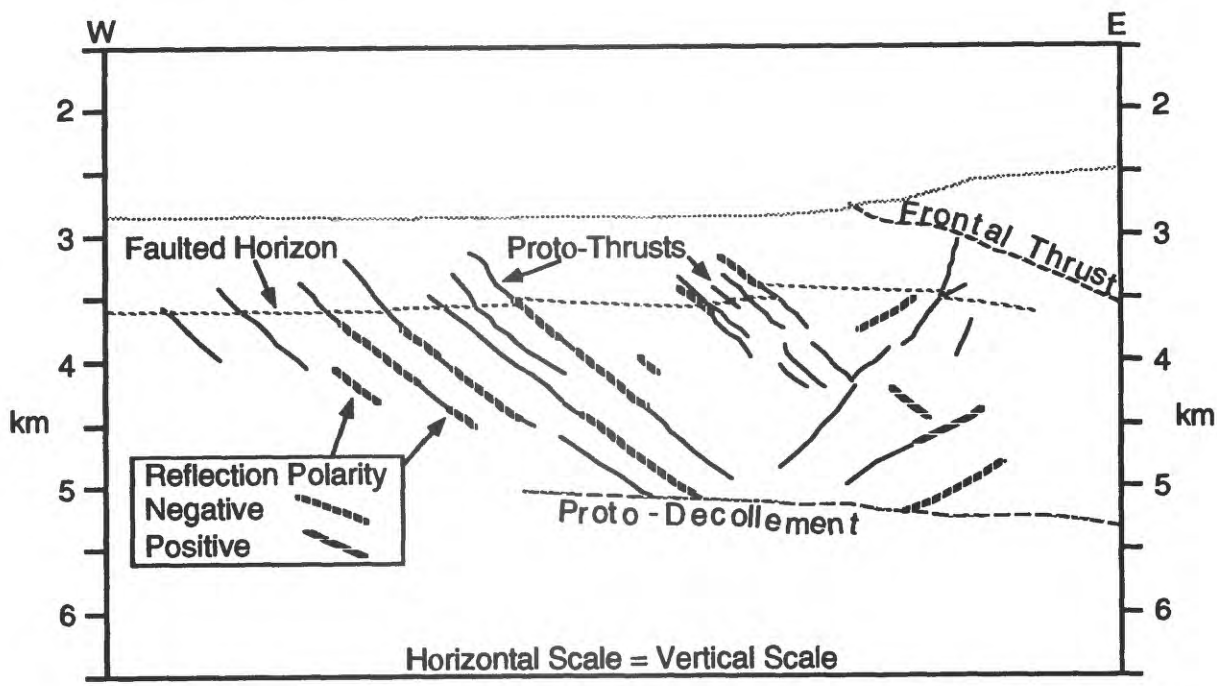



Figure 4. Cross section of seismic line 7 showing how a prominant horizon at about 750m depth was traced across the protothrust zone and used as an indicator of proto-thrust offset. Tracing this horizon across lines 5-9 provided the data for Figure 5.

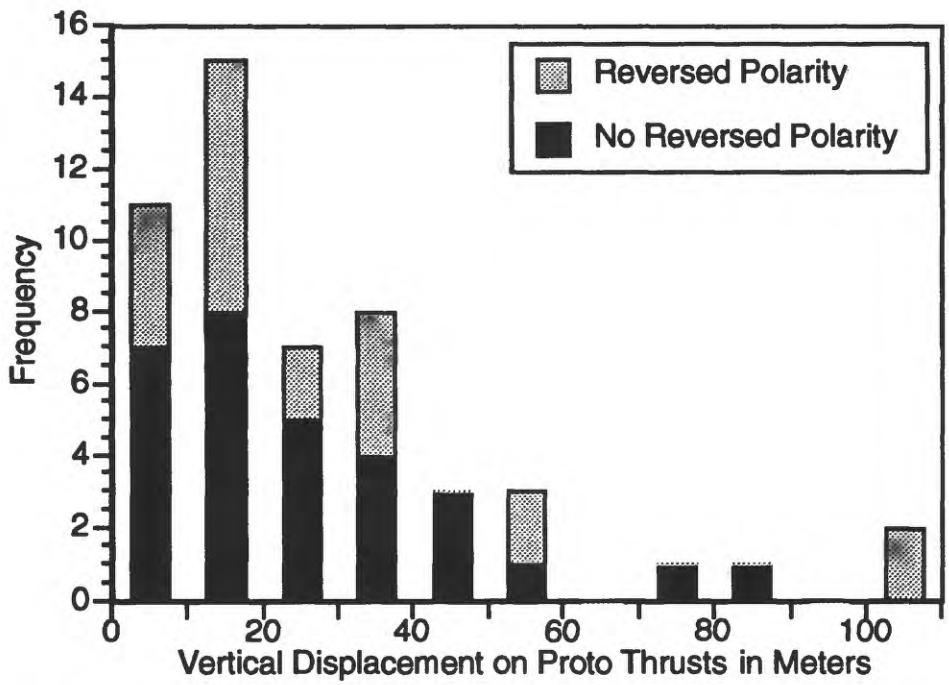



Figure 5. Histogram of offsets of proto-thrusts. Values for faults with and without reversed polarity reflections are separately tabulated. Note that reversed polarity reflections are not correlated with displacements.

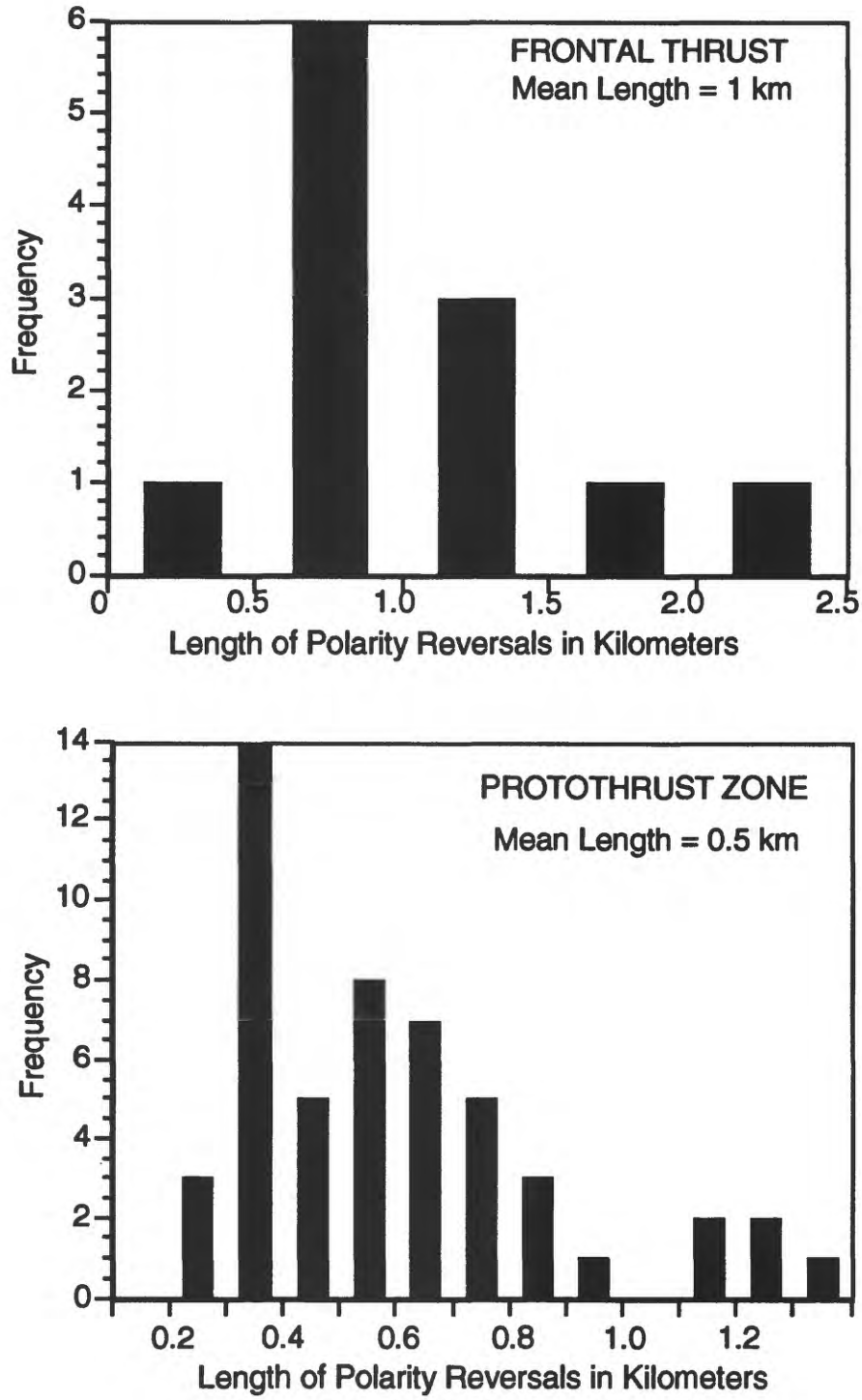



Figure 6. Histograms showing distribution of down-dip lengths of polarity reversals from both frontal thrust and proto thrusts.

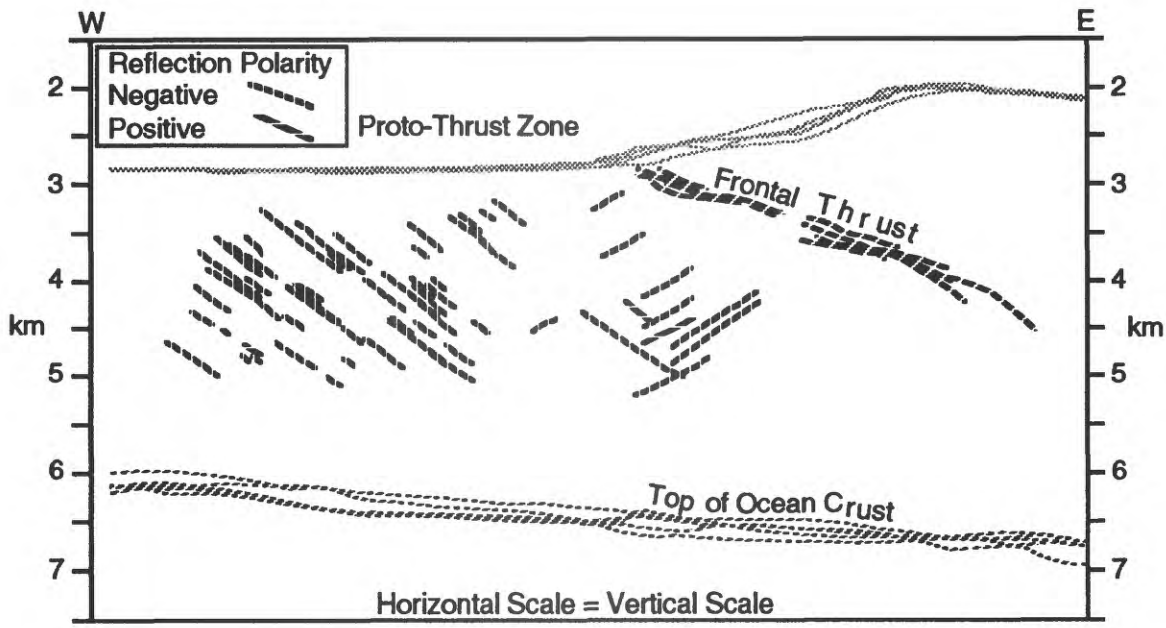



Figure 7. Geometry of polarity reversals on frontal thrust and proto-thrusts from lines 5-9 as viewed in composite E-W section (view south to north). Note segregation of polarity reversals on frontal thrust in to deep and shallow bands and lack of preferred depth distribution of polarity reversals on proto-thrusts.

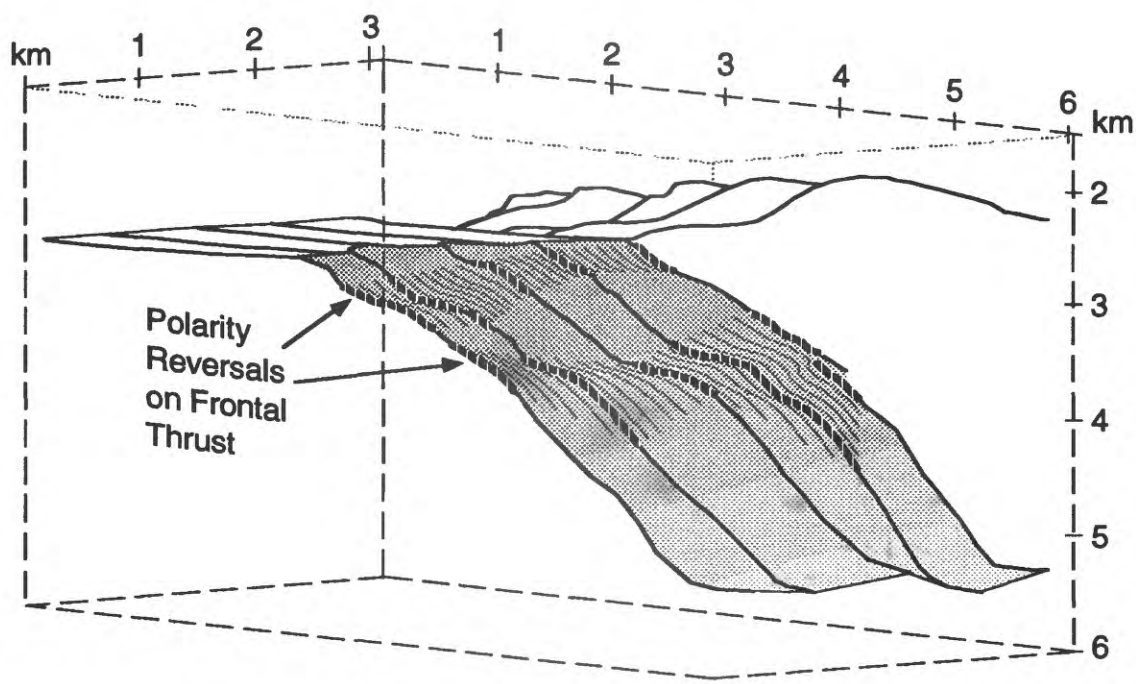



Figure 8. Perspective diagram viewing frontal thrust from southwest showing how polarity reversals occur in two bands extending across survey area.

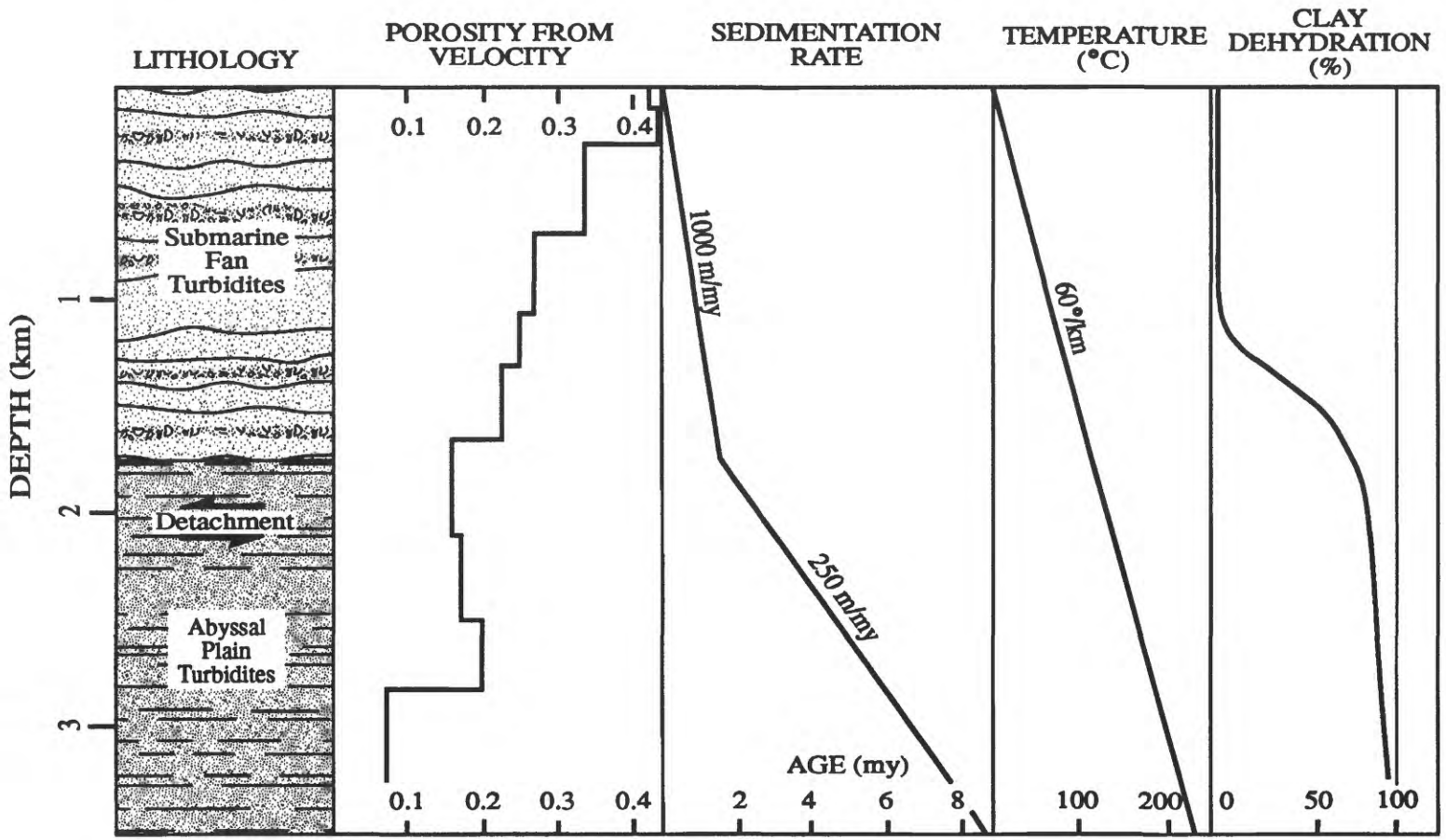



Figure 9. Conditions favoring high fluid pressues in the sedimentary sequence of the proto-thrust zone. Note the porosity inversion indicating underconsolidation and high fluid pressues at the base of the proto-thrust zone.

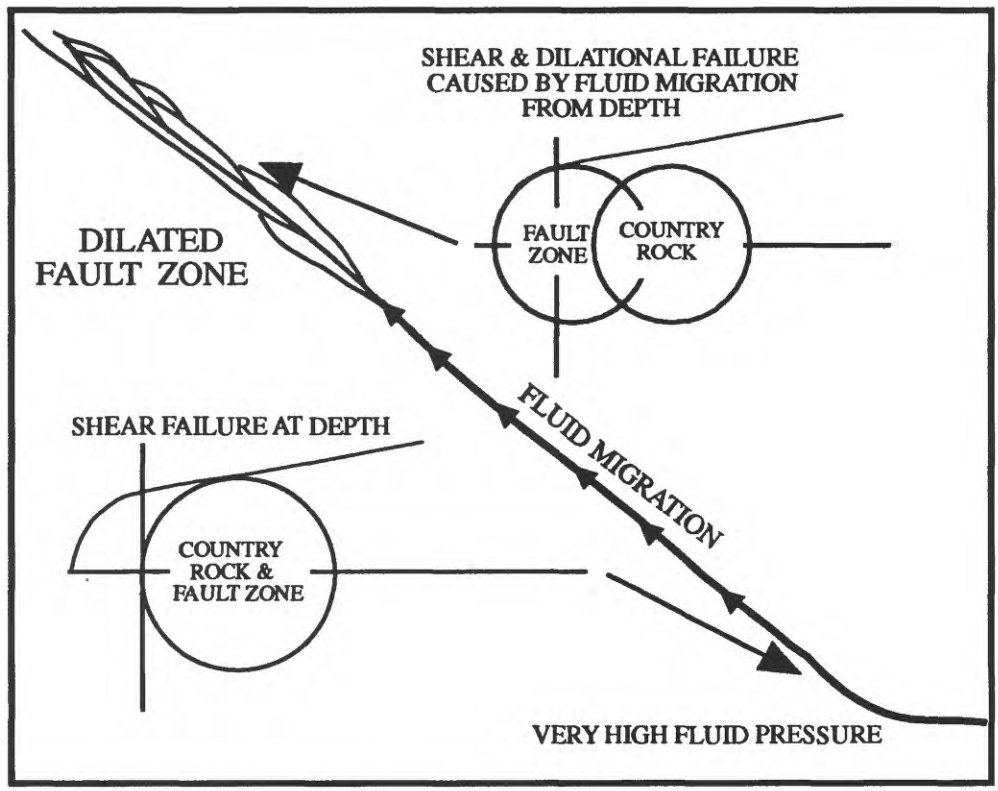



Figure 10. Geological-hydrological model of polarity reversals. Fluid is sourced in a high pressure zone and migrates up a steep fault. The drop in confining pressure exceeds head drop so that after migration fluid pressure in the fault zone exceeds the total pressure in the country rock and dilation occurs.

# ANATOMY OF A REGIONALLY EXTENSIVE FRACTURE NETWORK IN THE KODIAK ACCRETIONARY PRISM

Don M. Fisher, Susan L. Brantley, Mark Everett, Brett Wambold

Department of Geosciences, Pennsylvania State University, Univ. Pk, PA 16802

*Abstract*

Analysis of a system of quartz veins in the well-exposed central belt of the Kodiak Formation on Afognak Island, Alaska, allows documentation of a fracture network which served as the plumbing system of an underplated package of sediments near the base (8-12 km depth) of an ancient accretionary wedge. This vein system provides a record of high fluid pressure and hydrofracturing over a region at least 15 km across stike and 90 km along strike. The vein system displays a regular geometry, with thin, closely spaced (0.5-3 cm), near-vertical veins that connect vertically and/or laterally with thicker sigmoidal veins that are arranged in southeast-dipping *en echelon* sets. The thickness distribution for the veins is consistent with the equation, $N(X) \propto X^{-c}$, where N(X) is the number of veins with thickness > X and c = 1.3. Therefore, most of the total extension is accommodated by the thin veins: total extension is approximately 5%. Thin veins contain quartz with crack-seal textures; thicker veins contain inclusion-free quartz exhibiting euhedral terminations. Whole rock analysis of slates across the central belt and the structurally more shallow landward and seaward belts reveals significant variability in chemical composition, but little correlation between composition and structural level. Analysis of chemical composition on either side of four large euhedral growth veins reveals that silica is depleted on the side of the vein where quartz growth occurred. Pressure solution selvages trapped occasionally within the thicker veins and along their euhedral growth terminations contain high concentrations of immobile elements, indicating that these fractures closed by sudden collapse and penetration of quartz crystals into wallrock. A "zipper" model of fluid flow is consistent with textures and the fracture system geometry: lenses of fluid (<1m ) move upward through the thin veins as quartz, diffusing from wallrock into the open fractures, seals the base of the fractures. Thin fractures act as valves for fluid which collects in the thicker fractures aligned as shear zones; in these fractures, quartz growth, also driven by diffusive mass transfer, occurs until network interconnection or seismic event allows catastrophic fluid escape and fracture collapse. Fluid escaping through the network may be derived from local dewatering reactions as well as from deeper in the subduction zone.

## Introduction

*In situ* observations along active submarine accretionary prisms have recently led to many significant advances in our understanding of the hydrology of subduction zones. Rapid loss of porosity at the deformation front [*Bray and Karig* [1985]] attests to grain-scale fluid flow and expulsion of pore fluids. The existence of carbonate crusts and chimneys [*Kulm and Suess*, 1990; *Moore et al.*, 1990], thermal anomalies [*Foucher et al.*, 1990; *Fisher and Hounslow* , 1990], Cl anomalies [*Gieskes et al.*, 1990], mud volcanoes [*Brown and Westbrook*, 1988], and biological communities [*Lewis and Cochrane*, 1990; *Suess et al.*, 1985] along emergent thrusts at the toe of the accretionary wedge suggests that channelized fluid flow along faults must also occur. In some seismic reflection profiles, these faults display negative polarity anomalies [*Moore et al.*, 1990; *Moore and Vrolijk*, 1991] that may be an indication of fault-parallel dilated zones that move episodically [*Moore et al.*, 1991].

Much of the fluid that is expelled along faults may be derived from deeper levels, but our understanding of the deeper part of the forearc plumbing system is less complete and limited to studies of uplifted ancient accreted rocks (e.g., *Bebout* [1991]; *Vrolijk et al.* [1988]). In this paper, we examine the geometry, textural history, and wall rock geochemical variations associated with a regionally extensive vein system within the Kodiak Accretionary complex, an ancient subduction complex in southwest Alaska. The well-constrained tectonic setting and good exposure of the Kodiak Accretionary complex provides an exceptional opportunity to characterize the plumbing system of an underplated package of sediments near the base of an ancient accretionary wedge.

## Geologic Setting

The Kodiak accretionary complex consists of a series of northeast-trending belts of accreted rocks that decrease in age toward the Aleutian trench (Figure 1). The most areally extensive tectonostratigraphic unit within the exposed forearc is the early Maastrichtian Kodiak Formation, an 80 km-wide sequence of folded and imbricated turbidites that were accreted along the convergent margin in the Late Cretaceous or early Tertiary when there was a thick pile of sediments on the incoming plate [*Byrne and Fisher*, 1987; *Fisher and Byrne*, 1987]. The earliest recognized deformation within the Kodiak Formation resulted in stratal disruption in the footwall of the Uganik Thrust, where the Kodiak Formation is underthrust beneath the older Uyak Complex. $D_1$ also

resulted in the development of layer perpendicular calcite and quartz veins that are restricted to sand beds. $D_1$ structures, which formed when the sands and muds were unlithified or partially lithified, are interpreted to represent underthrusting in an early regime of progressive dewatering and lithification [*Fisher and Byrne*, 1987].

A later deformation event ($D_2$) involved folding, imbrication, slaty cleavage development, and lower greenschist facies metamorphism. This event occurred prior to the intrusion of 60 Ma plutons [*Sample and Moore*, 1987; *Fisher and Byrne*, 1992], indicating that $D_2$ structures formed during and shortly after accretion of the Kodiak Formation. Based on analyses of vitrinite reflectance, illite crystallinity, d-spacing of transitional graphite, and fluid inclusion analyses, this deformation and metamorphism must have ocurred at depths of 8-12 km [*Sample and Moore*, 1987; *Myers*, 1987; *Fisher and Brantley*, 1992], suggesting that the Kodiak Formation was underplated due to footwall collapse and duplex development along the plate boundary [*Sample and Fisher*, 1986]. The Kodiak Formation has been subdivided into the landward (LB), central (CB), and seaward belts (SB) (Figure 1) based on variations in $D_2$ structural style [*Sample and Moore*, 1983] .

A regional scale anticlinorium ($D_3$) folds the slaty cleavage and the imbricate thrust packages of the Kodiak Formation and exposes 1-5 km of $D_2$ structural relief from the core to the limbs (Figure 1). Landward and seaward belt rocks represent shallower structural levels while central belt rocks represent the deepest rocks exposed. From higher to lower levels, $D_2$ folds vary from tight and upright to isoclinal and recumbent, cleavage varies from vertical to subhorizontal, elongations measured from pressure shadows vary from 0.5 to 3.0, and strain histories vary from coaxial to noncoaxial [*Fisher and Byrne*, 1990; 1992]. In addition, there are pervasive $D_2$ veins at the lower structural level which are nearly absent in the highest exposed level [*Fisher and Byrne*, 1990]. Based on these considerations, we interpret the lowest $D_2$ structural level as a regional scale shear zone above a basal decollement that was active in the Late Cretaceous-Paleocene.

There are two regionally extensive sets of quartz veins associated with $D_2$. The earlier set ($D_{2a}$) is consistently oriented at about 70° clockwise relative to bedding in sand layers (viewed to the southwest) and at 15-35° clockwise to bedding in shale layers. Given regionally observed noncoaxial strain histories related to southeast-directed overthrusting on the decollement [*Fisher and Byrne*, 1992], the orientation of early $D_{2a}$

veins is best explained by 60° of counterclockwise angular shear in shale layers and 20° of angular shear in sand layers during slaty cleavage development.

Another vein set ($D_{2b}$) develops in the later stages of cleavage development. In an earlier paper, three types of $D_{2b}$ vein textures were described (Figure 2) [*Fisher and Brantley*, 1992]: (1) quartz veins with continuous bands of solid mica inclusions that record anywhere from 1 to 1000 episodes of cracking and sealing (similar to crack-seal textures described by *Cox and Etheridge* [1983], *Etheridge et al.* [1984], and *Ramsay* [1980]), (2) quartz veins with discontinuous crack-seal bands of solid mica inclusions that are restricted to quartz grains that widen toward the margin of the vein, and (3) inclusion-free quartz exhibiting euhedral terminations which, at the margin of the vein, are associated with insoluble residues in the matrix. We interpret these variations to indicate (1) periodic sealing of fractures (continuous crack-seal veins), (2) periodic partial sealing of fractures (discontinuous crack-seal veins), and (3) growth of quartz in an open crack with eventual collapse of the crack and dissolution of the matrix at impinging euhedral quartz terminations (euhedral growth veins). In our previous paper, we used textural arguments that silica transport involved local diffusion from wall rock to vein. However, we lacked information concerning the distribution of textures at the mesoscopic or regional scale.

This study focuses on the $D_{2b}$ veins in the lowest structural level of the Kodiak Formation. Three dimensional characterization of the $D_{2b}$ vein network is facilitated by near continuous cliff and wavecut platform exposures along the fjiord-indented northeast coastline of Afognak Island. The main features that distinguish $D_{2b}$ veins from $D_1$ veins in the field are that (1) the margins of most $D_{2b}$ veins are smooth at the grain scale and cut across grains in the wall rock rather than anastomosing around sand grains, (2) $D_{2b}$ veins are concentrated in shale-dominated units rather than in sandstones, (3) $D_{2b}$ veins cut across small scale sand-shale interlayers, and (4) $D_{2b}$ veins display evidence for continuous crack-seal (dark bands of solid inclusions parallel to the vein), discontinuous crack-seal (solid inclusion bands limited to isolated quartz grains), and euhedral growth (jagged margin on one side of vein) textures.

In this paper, we quantify vein spacing and thickness distributions to evaluate the mesoscale geometry of the fracture network. We then describe the distribution of the three observed textures within this system. Finally, variability in wall rock composition

around veins and throughout the Kodiak slate belt is determined to test our model for silica transport.

## Methods

### *Scanline measurements*

Measurement of orientation, thickness, length, and spacing of veins were made using scanline surveys [e.g.*Priest and Hudson*,1981]. Sixteen scanlines were measured on each of 16 outcrops in the central belt. Scanlines averaged 2.1 m in length, but varied up to 11 m depending upon extent of exposure. Each scanline consisted of a tape measure stretched over a planar or near planar rock surface. For every scanline, the length and orientation of the scanline, along with the position and thickness of every vein which crossed the line, were measured. Distance between veins (spacing) was measured from vein center to vein center. Vein orientations were measured along each scanline for a random fraction of veins. All vein thickness and spacing data was then corrected for scanline orientation using average vein orientation, and presented as true thickness and spacing.

### *Chemical composition analysis*

Slate samples from 11 localities were cleaned, cored, ashed to 900°C, and ground to -100 mesh. In all cases, samples were selected so as to avoid entraining vein material with slate matrix, and 2 to 4 samples were cored per location. Two of the 11 samples were also sampled for variability in chemical composition around a large euhedral growth vein. After grinding, samples were then reheated and dissolved using the lithium metaborate fusion method. Sulfur was determined using a LECO sulfur titrator; all other elements were determined with a Leeman Labs PS3000 inductively coupled plasma emission spectrometer (ICP). All results are presented on an ash-basis. Densities of two samples were measured using the immersion technique.

Polished thin sections of selected samples were also analyzed using a Cameca SX-50 electron microprobe. For compositional analyses on either side of quartz veins, intensity collections on samples and standards were completed using a focused beam

rastering over an area of 110 μm x 90 μm. Beam conditions for these analyses were 15.0 kV and 12 namps, and analyzer crystals were TAP, PET, and LIF crystals for Na, Al, and Mg; K and Si; and Ti, Fe, and Ca respectively.

*Microscopic measurements*

Thin sections of several samples were prepared and observed under optical microscope. Quartz band thickness was measured for all veins which showed continuous banding in thin sections from 5 localities. Measurements were made from the earliest crack-seal event (where quartz is fine-grained) to the latest event (at the margin(s) of the vein where quartz is coarse). The earliest event is commonly at the margin of the vein except in cases where cracking has occurred at both sides of the vein and the fine-grained quartz is observed in the vein interior. The angle between the vein margin and inclusion trails was also measured for each vein. Curved trails were broken into segments and the angle was measured for each segment.

## Results

*Outcrop scale observations of $D_{2b}$ veins*

The $D_{2b}$ vein set strikes northeast-southwest and occurs at a high angle to bedding in both argillite and sandstone layers. The $D_{2b}$ vein system displays a regular geometry, with thin, closely spaced (0.5-3 cm), near-vertical veins that connect vertically and/or laterally with thicker sigmoidal veins that are arranged in *en echelon* sets (Figures 3 and 4). Because the veins are constant in orientation, the distribution of veins can be easily evaluated along scanlines. Scanline surveys of argillites throughout the central belt reveal a log-normal distribution for the spacing of veins and spacing of *en echelon* sets with medians of 5.8 mm and 514 mm, respectively (Figure 5, see also, *Everett* [1993]). The spacing of fractures and brittle-ductile shear zones is typically very regular at the scale of tens of meters, and the overall geometry of the system ( thick *en echelon* veins within brittle-ductile shear zones and pervasive thin veins) is consistent throughout the central belt (i.e., 15 km of across-strike exposure). *Myers* [1987] also noted a predominance of *en echelon* sets of veins in the central belt between Kodiak and Afognak Islands, so this regular outcrop scale system extends at least 90 km along strike. The brittle ductile shear zones dip moderately to the southeast (trenchward) and contain sigmoidal veins which

indicate southeast-side down shear. There is a rarely observed conjugate *en echelon* set that dips northwest, and the undistorted tips of veins are parallel (i.e. vertical) in both northwest and southeast dipping zones.

*Chemical composition analysis*

Chemical composition of slates across the landward, seaward, and central belts showed significant variability in chemical composition, but little correlation between composition and structural level (Figure 6). Although the relationship between samples and structural level cannot be precisely quantified, the samples representing the deepest structural level lie approximately between kilometer 15 and 25. Averages of values of oxide wt. % for three samples each for the landward belt (LB), the deepest portion of the central belt (CB), and the seaward belt (SB) are summarized in Table I, along with one standard deviation ($\sigma$) around the mean. For each of three sample localities, 2 to 4 samples were analyzed in order to investigate variability across the transect as well as locally. In most cases, more variability in composition was observed in samples from the landward and seaward belts as compared to the central belt. Variability in wt. % oxide / wt. % oxide $Al_2O_3$, as documented by the standard deviation around the mean of this ratio for all samples, increased in the order $TiO_2 < Fe_2O_3 < SiO_2 < K_2O < MgO, Na_2O, CaO$.

**Table I. Chemical Composition of Landward, Central, and Seaward Belt Slate Samples (wt. %)**

| $SiO_2$ | $Al_2O_3$ | $TiO_2$ | $Fe_2O_3$ | MgO | CaO | MnO | $Na_2O$ | $K_2O$ | |
|---|---|---|---|---|---|---|---|---|---|
| 58.72 | 19.06 | 0.94 | 7.40 | 2.59 | 0.61 | 0.08 | 1.89 | 3.27 | Avg (LB) |
| 20.83 | 6.83 | 0.35 | 2.85 | 0.96 | 0.35 | 0.03 | 0.20 | 1.20 | $\sigma$ (LB) |
| 58.79 | 19.07 | 0.95 | 8.67 | 3.27 | 0.99 | 0.11 | 3.03 | 2.68 | Avg (CB) |
| 2.69 | 1.31 | 0.08 | 0.88 | 0.29 | 0.37 | 0.02 | 0.61 | 0.60 | $\sigma$ (CB) |
| 63.58 | 15.83 | 0.80 | 7.51 | 2.85 | 0.82 | 0.12 | 2.43 | 2.22 | Avg (SB) |
| 6.42 | 3.34 | 0.19 | 0.85 | 0.43 | 0.18 | 0.01 | 0.40 | 1.02 | $\sigma$ (SB) |

Table II and Figure 7a summarize chemical analyses of samples of slate matrix cored from four sites each from samples 3-85 and DC-2 (both from the central belt). Two samples were taken on each side of each vein, and the compositional analyses are

presented sequentially in Table II starting with samples of the wallrock on the side of the vein where euhedral growth terminations have formed (samples 1 and 2, not in the pressure solution selvage) followed by samples of the wallrock on the smooth side of the vein (samples 3 and 4). The ratio of $SiO_2$ to $Al_2O_3$ is lower on the growth side of the vein (samples 1 and 2). Variability in the ratios of $TiO_2/Al_2O_3$, $Fe_2O_3/Al_2O_3$, $MgO/Al_2O_3$, and $CaO/Al_2O_3$ were minor or absent between analyses (e.g.Figure 7a).

**Table II. Chemical Composition of Slate Samples around Veins (wt. %)**

| $SiO_2$ | $Al_2O_3$ | $TiO_2$ | $Fe_2O_3$ | MgO | CaO | MnO | $Na_2O$ | $K_2O$ | Locality | # |
|---|---|---|---|---|---|---|---|---|---|---|
| 56.6 | 20.9 | 1.12 | 9.90 | 3.80 | 1.35 | 0.13 | 3.29 | 2.91 | 3-85A | 1 |
| 56.7 | 20.7 | 1.06 | 8.66 | 3.39 | 1.53 | 0.11 | 3.01 | 3.21 | 3-85A | 2 |
| 63.5 | 17.7 | 0.91 | 8.16 | 3.08 | 0.96 | 0.09 | 2.99 | 2.46 | 3-85A | 3 |
| 60.7 | 18.2 | 0.94 | 8.25 | 3.15 | 1.61 | 0.13 | 2.74 | 2.61 | 3-85A | 4 |
| | | | | | | | | | | |
| 53.8 | 18.9 | 0.88 | 9.62 | 3.23 | 0.65 | 0.11 | 3.57 | 2.32 | DC2 | 1 |
| 54.6 | 15.4 | 0.75 | 9.56 | 3.07 | 0.61 | 0.11 | 3.14 | 1.54 | DC2 | 2 |
| 61.2 | 17.2 | 0.83 | 8.78 | 2.92 | 0.67 | 0.11 | 3.41 | 2.04 | DC2 | 3 |
| 61.6 | 16.0 | 0.77 | 8.62 | 2.78 | 0.63 | 0.1 | 3.29 | 1.85 | DC2 | 4 |

Microprobe analysis of whole rock composition on either side of two euhedral growth veins (samples DC-6 and BCK, both from the central belt) also documented that the matrix next to the side of the vein with crystal terminations (the growth side) was depleted in silica compared to the flat side of the vein. The profile of $SiO_2/Al_2O_3$ for sample DC-6 and BCK show a strong depletion in silica around the vein (Figure 7b,c). However, silica depletion around a continuously banded crack-seal vein in a sample from site DC-6 (total thickness ≈ 0.3 mm) was unmeasureable: $SiO_2/Al_2O_3 = 3.41 \pm 0.16$ on the growth side of the vein and $3.40 \pm 0.21$ on the initial side of the vein.

Microprobe analysis also revealed that "mica" grains in the matrix of sample 3-85 are an illite with composition $(Na_{0.77}K_{0.38})(Al_{2.6}Mg_{0.82}Fe_{1.06})Si_{7.29}Al_{0.71}(O_{10}(OH)_2)$. Very little variation in composition between analyses was observed. Phyllosilicate grains that are included within the quartz veins are chlorite in composition, with an average

Mg/Fe ratio of 0.74. Chlorite of the same composition has also been identified to be common in the wallrock. Variation in chlorite composition is approximately 9% in the Fe/(Fe + Mg) ratio.

Microprobe analysis documented the presence of enriched phosphorus, titanium, and rare earth elements in the pressure solution selvage observed at the edge of the euhedral growth vein of sample 3-85. Based on these analyses, the phases within the selvage include 20-μ-sized grains of rutile, apatite, and monazite.

*Textural Observations*

The microscale textures of $D_{2b}$ veins vary with respect to the systematic fracture network. For example, the thin veins between *en echelon* sets are continuous crack-seal veins with 1-50 planar, continuous bands of chlorite (with occasional other phase) inclusions. Measured thickness of quartz bands reveals a lognormal distribution with a mean and median of 9.0 and 7.1 μm respectively (Figure 8). The veins within *en echelon* sets show all three types of textures but are dominated by discontinuous crack-seal and euhedral growth textures. Continuously banded crack-seal veins in the shear zones were observed to have as many as 1000 cracking events.

Measurements of the thickness of crack-seal events, made sequentially from the side of the vein where fracturing initiated (quartz is fine-grained) to the side of the vein where the latest event occurred (where quartz is coarse-grained), revealed that in 13 of 15 veins studied, band thickness decreased over time (e.g. Figure 9). Figure 9 also shows that the trend of decreasing thickness as a function of time is observed in three places along the same crack-seal vein, although the exact thickness at any given point for any given band varies.

Inclusion trails within veins provide a record of the displacement associated with cracking [*Cox and Etheridge*, 1983; *Cox*, 1988]. In thin veins between *en echelon* sets, the inclusion trails are typically perpendicular to the margin of the vein, indicating that cracks experience mode I failure. Veins within *en echelon* sets initiate as mode I cracks but along the most recently active side of the vein, the inclusion trails are commonly oblique to the vein margin.

The thick veins also display textural evidence for periodic collapse of the fluid filled pockets. Dark insoluble residues occur in the wallrock at the margins of euhedral growth veins where collapse of the wall rock around euhedral quartz crystal terminations resulted in dissolution of silica. Jagged insoluble residues that are embedded within the quartz vein record earlier crack collapse events. The spacing of these features within veins varies from 0.22 to 11.1 mm.

## Discussion

*Regional distribution of veins*

The $D_{2b}$ veins accommodate subhorizontal extension perpendicular to the trace of the present-day plate boundary. The undistorted tips of veins in the southeast-dipping and rare northwest-dipping shear zones are parallel, indicating that vein arrays develop as a consequence of initiation and propagation of vertical fractures, and that shearing deformation is subsequently localized along zones of low competence related to the presence of the fractures (e.g., type II *en echelon* sets, [*Beach*, 1975]). The dominant *en echelon* set (southeast-dipping zones) is favored because the sense of slip on these zones is synthetic with the sense of slip along the underlying plate boundary.

One method of evaluating the relative importance of thin and thick veins for the total extension is to examine the size distribution of veins. The size distribution of earthquakes and fault populations has been shown to obey the relation:

$$N(X) = a\, X^{-c} \qquad (1)$$

where N(X) is the number of earthquakes of seismic moment > X or the number of faults of length or displacement > X [*Scholz and Cowie,* 1990; *Marrett and Allmendinger*, 1992]. The $D_{2b}$ veins in this study are analogous to faults in the sense that the total displacement on a fault represents the cumulative result of many repeated slip events and the total thickness of $D_{2b}$ veins represents the cumulative effect of numerous crack-seal events. The thickness distribution of $D_{2b}$ veins in the central belt is consistent with equation 1, where N(X) equals the number of veins with thickness > X, for a value of c = 1.3 (Figure 10). In the field, any vein with thickness less than 1 mm was recorded as a "thin" vein; measurement of 250 of these "thin" veins under the microscope revealed a lognormal distribution with a mean of 0.2 mm. Therefore, in Figure 10, the data point for "thin" veins is plotted at X = 0.2 mm. However, the regression was calculated only

through data points where X > 1 mm. Because vein thickness is well-described by equation (1), it follows that vein thickness displays a fractal distribution with fractal dimension of 1.3 [*Turcotte,* 1986]. Based on these results, most of the total extension related to $D_{2b}$ veining is mostly accommodated by the pervasive closely spaced thin veins (e.g. *Scholz and Cowie*, 1990). Total elongation, calculated as vein thickness / (vein spacing - vein thickness), averages about 5% in the slates.

*Compositional analysis*

Figure 6 documents that local variability in oxide components is larger than any trends which exist across the slate belt. Lack of variability in the $TiO_2/Al_2O_3$ ratio (standard deviation around the mean for the entire belt = 6%) suggests a lack of variability in those components in the original sediments and that those components are immobile. Variability in Fe and Mg distribution may be related to variability in precipitation of chlorite, pyrite, and iron oxides throughout the slates, as suggested for the Martinsburg Fm. by *Wintsch et al.* (1993). The largest variability (standard deviation = 36 to 41%) was observed for $Na_2O/Al_2O_3$ and $CaO/Al_2O_3$ ratios across the slate belt (Figure 6). Na and Ca were also observed by *Wintsch et al.* (1993) to show a large variability within the Martinsburg slate belt, which those authors attributed to mobility on the scale of kilometers.

*Wintsch et al.* (1993) concluded that $SiO_2$ mobility was limited to the scale of centimeters, and that mudstones showed no loss of silica that could be correlated with strain throughout the Martinsburg. The absence of a trend in the $SiO_2/Al_2O_3$ ratio as a function of position in the Kodiak slates (Figure 6) may also be evidence that no long-distance transport of silica has occurred. However, the standard deviation around the mean of the $SiO_2/Al_2O_3$ ratio for the entire slate belt is $\pm$ 27%, which is much larger than that observed for $TiO_2$ ($\pm$6%). Local mobility of silica is large.

If vein silica is derived from wallrock, as suggested by our model, we might expect that matrix composition for central belt samples with veins would be lower in silica than comparable matrix slate samples from the seaward and landward belts, both of which lack $D_{2b}$ veins. The average of the ratio $SiO_2/Al_2O_3$ in the deepest samples of the central belt ($3.1 \pm 0.4$) was lower than the seaward belt average ($4.0 \pm 1.4$), but not lower than the landward belt average ($3.1 \pm 1.1$). Much of the variability in this ratio must not be related to the growth of quartz veins.

According to our model of silica diffusion from wallrock to vein, silica should be more depleted in the wallrock adjacent to the open fracture. The analyses of four veins indicate that silica depletion (and enrichment of other components) has occurred on the active side of the vein (Table II, Figure 7). Results are comparable for samples measured using whole rock chemical analysis (Table II, Figure 7a) and microprobe analysis (Figure 7b,c). The only vein studied which did not show silica depletion on the growth side was a 1-mm wide crack-seal vein. These results are therefore in agreeement with a model of diffusive mass transport of silica from matrix to vein. Silica mobility is present at least at the centimeter scale, while Fe, Ti, Al, Mg, Mn and other components show no variability with respect to position around quartz veins.

If aluminum is immobile in these samples, then the volumetric strain of the depleted side of the vein with respect to the undepleted side of the vein must equal $(C_u\rho_u - C_d\rho_d)/C_d\rho_d$, where the wt. % $Al_2O_3$ and the density of the depleted zone are $C_d$ and $\rho_d$ respectively and the wt. % $Al_2O_3$ and density of the undepleted side are $C_u$ and $\rho_u$ respectively. The calculated volumetric strain ($\Delta V/V_o$, where $\Delta V$ is the change in volume and $V_o$ is the original volume) for samples 3-85 and DC-2 are -23% and -18% respectively, assuming that $\rho_d \sim \rho_u$. For the one vein where we measured density on either side of the vein (DC-2), this assumption proved true ($\rho_d$ = 2.744 g/cm$^3$; $\rho_u$ = 2.748 g/cm$^3$). This calculated strain is a minimum estimate, and represents the strain due to mobilization of silica assuming that one side of the vein has maintained its original silica content. The elongation represented by vein material within the shear zones, as calculated from scanlines, varies between about 8 and 30%. Loss of silica from the wallrock adjacent to the vein corroborates the model of silica diffusion into the open fracture; for the few measurements completed, enough silica has been mobilized to account for the elongation recorded by veins. The only vein which did not show silica depletion was a 1-mm wide crack-seal vein. Resolution of our analysis would not determine silica depletion on the order of 5-6 %, suggesting that the strain calculated for this sample was < 6%. The elongation associated with the crack-seal veins (outside of the shear zones) is approximately 5%, as determined from scanline measurements, and therefore the depletion zone around those veins would not be detectable by the microprobe measurements.

*Interpretation of crack-seal vein textures*

Each band represents a crack-seal event that involves cracking along the boundary between the quartz vein and the wallrock, followed by sealing of the crack due to growth of chlorite (predominantly) from wall rock seed crystals and growth of quartz from the surface of the vein (Figure 2). Given the crystallographically controlled anisotropy in quartz growth rate within open fractures, crack collapse prior to complete sealing results in preservation of the roughness of the vein surface at the time of collapse [*Fisher and Brantley*, 1992]. The continuity and planarity of inclusion bands in thin veins suggests that cracks are completely sealed with quartz along their length. Under these circumstances, the thickness of inclusion bands is a measure of the crack aperture at the time of each sealing event, or at least, if fractures open widely and then relax, a measure of the minimum crack aperture. Based on microscopic measurements of inclusion band spacing in thin veins, the apertures associated with cracking events show a log-normal distribution with a mean and median of 9.0 and 7.1 μm respectively (Figure 8).

Most $D_{2b}$ veins within the Kodiak Formation develop asymmetrically with all or most of the cracking on one side of the vein. Of all the veins observed, only 19 of 239 showed growth on both sides of the veins. For the thin continuous crack-seal veins between *en echelon* sets, there is a general decrease in the spacing of inclusion bands, suggesting that the crack aperture associated with each cracking event decreases over time (Figure 9). Continuous crack-seal bands consistently exhibit inclusion trails aligned perpendicular to vein walls, so a decrease in vein aperture cannot be explained by mixed mode cracking with a constant displacement. The aperture of a mode I crack is determined by a number of variables such as the fluid pressure within the crack and the elastic properties of the material. For example, Young's modulus (E) could progressively increase in the rock due to the introduction of quartz within veins. An increase in stiffness would result in a smaller crack aperture for a given crack length. However, the amount of quartz associated with $D_{2b}$ veins (approximately 5 volume %) would not produce sufficient change in E to account for the significant decrease in aperture observed. Another explanation for the decrease in crack aperture over time is that the fracture toughness of the rock decreases due to weaknesses or the presence of flaws along the boundary between quartz veins and the wall rock. Under these circumstances a smaller increase in fluid pressure is required for critical crack growth.

If the initial *in situ* aperture of the crack is larger than the thickness of the quartz band, then the band thickness may also be a function of the balance between the rate of quartz precipitation vs. the rate of relaxation of the fracture aperture. We know that the

rate of precipitation is greater than the rate of relaxation because no crystal terminations are recorded in the crack-seal veins--fracture closure occurs by sealing rather than by collapse. A decrease in band thickness could therefore also indicate an increase in the rate of quartz precipitation over time, assuming a constant rate of aperture relaxation.

*Interpretation of shear zone vein textures*

Individual veins within the shear zones typically originate as continuous crack-seal veins but evolve into discontinuous crack-seal veins and finally euhedral growth veins. This is consistent with our model for formation of the *en echelon* zones: distributed vertical hydrofracturing leads to zones of weakness that evolve into shear zones. Once brittle-ductile shear zones nucleate, the displacement associated with cracking episodes increases for veins within the zone and *en echelon* cracks persist as open fluid-filled reservoirs. The observation that phyllosilicate inclusion trails trapped in the shear zone veins are commonly oblique to vein walls in later stages of formation documents that there is a component of displacement parallel to the crack during the later stages of vein formation in those zones. The relative displacement of opposing sides of the crack is antithetic to the sense of shear within the southeast dipping brittle-ductile shear zones. The increasing component of slip parallel to the cracks is consistent with the observed clockwise rotation of *en echelon* veins relative to the stress field. No consistent relationship was observed between the crack aperture (thickness between inclusion bands) and the opening direction (orientation of inclusion trails).

The thick veins also display textural evidence for periodic collapse of the fluid-filled pockets. Dark insoluble residues occur in the wallrock at the margins of euhedral growth veins where collapse of the wall rock around euhedral quartz crystal terminations resulted in dissolution of silica. Jagged insoluble residues that are embedded within the quartz vein record earlier crack collapse events. The spacing of these features within veins varies from 0.22 to 11.1 mm. Thus, collapse events in *en echelon* veins are much less frequent than the more closely spaced crack-seal events. Collapse events are interpreted as evidence for periodic linkage of *en echelon* reservoirs and draining of fluid pockets. Phases documented in the pressure solution selvage, including rutile, monazite, and apatite, contain components documented to be immobile (e.g. *Wintsch et al.* [1991], *Ayers and Watson* [1991]).

## Models of Vein Formation and Fluid Flow

The mesoscopic distribution and textural variability of the $D_{2b}$ vein network suggests that the development of the vein system involves an early history of distributed hydrofracture nucleation throughout the lowest structural level of the Kodiak Formation, with an increasing density of veins as fractures seal and new fractures develop. This early history is followed by nucleation of southeast dipping brittle-ductile shear zones as the fracture density increases until arrays of fractures locally provide zones of weakness that experience shear failure. These *en echelon* cracks no longer seal completely but rather provide small reservoirs that remain open or only seal partially. The $D_{2b}$ vein system provides a record of high fluid pressure and hydrofracturing over an exposed region 15 km across strike and 90 km along strike.

In our previous paper, we developed a model for fluid flow that was consistent with observations of continuous crack-seal veins. We argued that, because the difference ($\Delta P$) between fluid pressure in the matrix and in the crack increases with depth within a fracture or interconnected fracture network, matrix-to crack diffusion and vein growth will be fastest at the lower tip of the fracture. Thus, vertical fractures will seal at depth while the open crack lens of fluid will propagate upward maintaining a steady state crack volume. Based on this model, the upward migration of fluids and cracks is rate-controlled by quartz growth at the base of the fractures and fluid lenses move up in a zipperlike fashion.

This "zipper" model can now be placed in the context of the observed mesoscale vein distribution. Lenses of fluid moving upward through the thin veins must be limited to vertical dimensions < 1 m (Figure 3). Such a dimension is consistent with the model of pore-fluid diffusive transport of silica into the vein which requires fluid lenses with vertical dimensions less than 6 m [*Fisher and Brantley*, 1992]. Because the textures of euhedral growth veins document that these fractures remained open for extended time periods while the textures of crack-seal veins document that these fractures only remain open intermittently, the thin veins must act as valves between shear-zone fluid reservoirs. Small fluid lenses migrate upward from reservoir to reservoir as quartz diffuses from the wall rock into the thin cracks.

Thin veins with as few as 1 and as many as 50 events have been observed, implying that the entire vein system cannot experience failure at the same time. Furthermore, where thicker veins in the shear zones have continuously banded crack-seal textures, up to 1000 bands have been observed. Therefore, distributed fluid lenses must all migrate between reservoirs independently. One-event crack-seal veins are always observed to have a vertical scale of < 2 cm, implying that fractures initiate as features which are small in length and which link into the network over time. The hundreds of crack-seal events recorded by individual quartz grains within discontinuous crack-seal veins may represent periodic inflation of the fluid reservoirs in shear zones due to the addition of each single lens of fluid. Collapse events in *en echelon* zones represent less frequent events when adjacent reservoirs become interconnected and are able to drain. Such collapse events may be caused by seismic events.

Measurement of silica depletion zones around four euhedral growth veins documents local mass transfer of silica from wallrock to fracture for those samples. *Fisher and Brantley* [1992] argued from textural observations for such a diffusive mass transfer model. A single fluid pulse model, whereby each quartz layer precipitates after influx of a single crack volume of fluid [*Etheridge et al.*, 1984], was ruled out because this model would require unreasonably large crack apertures to precipitate 8 μm of quartz [*Fisher and Brantley*, 1992]. A fracture-channelized flow model, whereby silica is derived from deeper levels and is precipitated during long distance flow down a temperature and pressure gradient [e.g. *Yardley*, 1984], is unlikely based on the observation that continuous crack-seal veins seal rather than collapse [*Fisher and Brantley*, 1992]. During sealing, fluid flow would become progressively slower until the rate of delivery of $SiO_2$ to the system would be effectively zero. Moreover, quartz growth in euhedral growth veins is greatest adjacent to sandy laminations in the wall rock, suggesting that wall rock composition influences axial concentration gradients within the crack. A diffusion model from wall rock to vein is favored because this model does not require enormous crack apertures and can result in complete sealing of fractures. This model allows for local derivation of silica and explains the silica depletion zones that have been identified in the wall rock adjacent to the open side of euhedral growth veins.

Although silica is derived from local wallrock in this model, fluid filling the fractures may be derived locally from dehydration reactions, or externally from deeper in the subduction zone. The fracture network documents only that a source of fluid was

available, and that this fluid left the system as small lenses pushed upward by quartz precipitation.

*Fisher and Brantley* [1992] postulated that reasonable rates of sealing could be obtained with pore-fluid diffusion from matrix to crack with a pressure differential on the order of 0.01 to 0.1 MPa. However, newly published data for diffusion along grain boundaries in novaculite in the presence of a wetting pore fluid suggests that grain-boundary diffusion rates in the presence of certain fluids may be orders of magnitude faster than previously hypothesized [*Farver and Yund*, 1992]. These results are consistent with inferences made by *Fisher and Brantley* [1992] that the rate of pressure shadow development in Kodiak slates also required grain-boundary diffusion coefficients that were faster than previously hypothesized. The new data for grain-boundary diffusion coefficients may be consistent with grain-boundary diffusion as the controlling mechanism of silica transport from matrix to vein.

*Fisher and Brantley* [1992] argued that sealing of the thin veins must be interface-controlled early in vein growth, and mixed interface- and diffusion-controlled late in vein growth. This implies that the rate of sealing of the thin veins increased with time. If the fracture aperture was larger upon initial fracturing and subsequently relaxed, then as the rate of sealing increased the thickness of precipitated quartz bands would have decreased over time, as observed. The model presented here, where thin veins are valves allowing fluid flow between shear zones, also implies that, if the rate of introduction of fluid is constant while the rate of sealing of the valves increases, then the rate of fracturing as documented in the crack-seal textures should also increase with time. In fact, analysis of the textures led *Fisher and Brantley* [1992] to conclude that the period between fracturing decreased during the growth of these veins.

## Conclusions

The $D_{2b}$ vein network exposed along the northeast coast of Afognak Island, Alaska provides a record of high fluid pressure and hydrofracturing within an ancient forearc plumbing system that extends for at least 15 km across strike and 90 km along strike. This network consists of two parts: (1) thin (<1mm) vertical veins that strike parallel to the margin and display a log-normal spacing distribution and a median spacing of 5.8 mm, and (2) thicker veins that are arrayed in en echelon sets that define shear zones that strike parallel to the margin and dip southeast with southeast-side-down shear. These

shear zones display a log-normal spacing distribution with a median spacing of 514 mm. Throughout the network, the veins display a power law thickness distribution with a fractal dimension of 1.3, indicating that thin veins accommodate the largest proportion of the bulk extension during $D_{2b}$ (approximately 5%).

Microtextures of veins within the fracture network indicate that thin veins experience periodic fracturing followed by complete sealing along the length of fractures. The spacing of inclusion bands in these veins provides a measure of minimum in situ crack aperture, with a log-normal distribution of crack seal band thickness and a median thickness of 7.1 μm. The thick veins within the shear zones record partial sealing of fractures or growth of quartz into open voids with intermittent collapse of cracks and dissolution of wall rock against quartz crystal terminations. Thus, the thick veins in brittle-ductile shear zones serve as reservoirs for fluids while thin veins act as valves that allow migration of high aspect ratio fluid lenses between shear zones. The inferences about thin veins are consistent with a zipper model for migration of fluid lenses [*Fisher and Brantley* 1992] where quartz precipitates at the base of cracks as fractures propagate upward. The collapse events in thick veins reflect catastrophic events where fluid pockets become interconnected and higher pressure reservoirs are allowed to drain into areas of lower pressure. The interconnection of fluid reservoirs is one of the possible ways in which seismic events may influence fluid migration around faults [e.g. *Byerlee*, 1993], so spaced collapse features may be related to paleoseismicity.

There is no evidence at present for long distance transport of silica within the Kodiak Formation based on regional variations in the chemical composition of slate samples. There is, however, geochemical evidence for local migration of silica from the rock matrix to veins. For thick veins that record quartz growth on only one side of the vein, there is a measurable depletion of silica in the wall rock on the active side of the vein relative to the side of the vein that did not have access to the open fracture. Observations of silica depletion zones are in agreement with inferences made based on textural and geochemical arguments for these veins [*Fisher and Brantley*, 1992]. The amount of wall rock depletion is comparable to the amount of quartz precipitated by veins. Thus, the $D_{2b}$ vein network provided distributed sinks for silica that was transported by diffusion from the wall rock to cracks during upward fluid migration along a fracture-dominated plumbing system.

## References


Ayers, J. C. and E. B. Watson, 1991, Solubility of apatite, monazite, zircon, and rutile in supercritical aqueous fluids with implications for subduction zone geochemistry, *Phil. Trans. R. Soc. Lond. A, 235*, 365-375.

Beach, A., 1975, The geometry of en-echelon vein arrays, *Tectonophysics, 28*, 245-263.

Bebout, G., 1991, Geometry and mechanisms of fluid flow at 15 to 45 km depths in an Early Cretaceous accretionary complex, *Geophys. Res. Lett., 18*, 923-926.

Bray, C. J., and D. E. Karig, 1985, Porosity of sediments in accretionary prisms and some implications for dewatering processes, *J. Geophys. Res., 90*, 768-778.

Brown, K. M. and G. K. Westbrook, 1988, Mud diapirism and subcretion in the Barbados Ridge accretionary complex: the role of fluids in accretionary processes, *Tectonics, 7*, 613-640.

Byerlee, J., 1993, Model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology, 21,* 303-306.

Byrne, T. and D. Fisher, 1987, Episodic growth of the Kodiak convergent margin, *Nature, 325*, 338-341.

Cox, S. F. and M. A. Etheridge, 1983, Crack-seal fiber growth mechanisms and their significance in the development of oriented layer silicate microstructures, *Tectonophysics, 92*, 147-170.

Etheridge, M. A., Wall, V. J., Cox, S. F., and R. H. Vernon, 1984, High fluid pressures during regional metamorphism and deformation, *J. Geophys. Res., 89*, 4344-4358.

Everett, M., 1993, The fracture network of the central belt of the Kodiak Formation, Alaska, Unpublished Masters Thesis, Department of Geosciences, Pennsylvania State University, University Park, PA.

Farver, J.R. and R.A. Yund, Oxygen diffusion in a fine-grained quartz aggregate with wetted and nonwetted microstructures, *J. Geophys. Research, 97*, 14031-14054, 1992.

Fisher, A. T., and M. Hounslow, 1990, Transient fluid flow through the toe of the Barbados accretionary complex: constraints from Ocean Drilling Program Leg 110 heat flow studies and simple models, *J. Geophys. Res., 95*, 8845-8858.

Fisher, Donald M. and Brantley, Susan L., Models of quartz overgrowth and vein formation: Deformation and episodic fluid flow in an ancient subduction zone, *J. Geophysical Research, 97,* 20,043-20,061, 1992.

Fisher, D. and T. Byrne, 1990, The character and distribution of mineralized fractures in the Kodiak Formation, Alaska: implications for fluid flow in an underthrust sequence, *J. Geophys. Res., 95,* 9069-9080.

Fisher, D. and T. Byrne, 1987, Structural evolution of underthrusted sediments, *Tectonics, 6,* 775-793.

Fisher, D. and T. Byrne, 1992, Strain variations in an ancient accretionary wedge: implications for forearc evolution, *Tectonics, 11,* 330-347.

Foucher, J. P., Le Pichon, X., Lallement, S., Hobart, M., Henry, P., Beneditti, M., Westbrook, G.K., and M. A. Langseth, 1990, Heat flow, tectonics, and fluid circulation at the toe of the Barbados Ridge accretionary prism, *J. Geophys. Res., 95,* 8859-8867.

Gieskes, J. M., Vrolijk, P., and G. Blanc, 1990, Hydrogeochemistry of the northern Barbados accretionary complex transect: ocean drilling project leg 110, *J. Geophys. Res., 95,* 8809-8818.

Kulm, L. D. and E. Suess, relationship between carbonate deposits and fluid venting: Oregon accretionary prism, *J. Geophys. Res., 95,* 8899-8916.

Lewis, B. T. R. and G. C. Cochrane, 1990, Relationship between the location of chemosynthetic benthic communities and geologic structure on the Cascadia subduction zone, *J. Geophys. Res., 95,* 8783-8794.

Marrett, R. and Allmendinger, R.W., 1992, Amount of extension on "small" faults: An example from the Viking graben, *Geology, 20,* 47-50.

Moore, G. F., Shipley, T. H., Stoffa, P. L., Karig, D. E., Taira, A., Kuramoto, S., Tokuyama, H., and K. Suyehiro, 1990, Structure of the Nankai trough accretionary zone from multichannel seismic reflection data, *J. Geophys. Res., 95,* 8753-8766.

Moore, J. C. and P. Vrolijk, 1992, Fluids in accretionary prisms, *Reviews of Geophysics, 30,* 113-135.

Moore, J. C., Cochrane, G., Mackay, M., Moore, G., Soest and K. Brown, 1991, Dilational waves along thrust faults of the modern Oregon accretionary prism, *Geol. Soc. Am. Abstr. with Prog., 23,* no. 5, A366. A355?

Moore, J. C., Orange, D., and L. D. Kulm, 1990, Interrelationship of fluid venting and structural evolution, Alvin observations from the frontal accretionary prism, Oregon, *J. Geophys. Res., 95*, 8795-8808.

Myers, G., 1987, Fluid expulsion during the underplating of the Kodiak Formation: a fluid inclusion study, Masters thesis, University of California, Santa Cruz.

Priest, S. D., and Hudson, J. A. Estimation of Discontinuity Spacing and Trace Length Using Scanline Surveys, *Int. J. Rock Mech. Min. Sci. & Geomech. Abstr.,18*, pp. 183-197, 1981.

Ramsay, J., 1980, The crack-seal mechanism of rock deformation, *Nature, 284*, 135-139.

Sample, J., and D. Fisher, 1986, Duplex accretion and underplating in an ancient accretionary complex, Kodiak Islands, Alaska, *Geology, 14*, 160-163.

Sample, J., and J. C. Moore, 1987, Structural style and kinematics of an underplated slate belt, Kodiak and adjacent islands, Alaska, *Geol. Soc. Am. Bull., 99*, 7-20.

Scholz, C.H. and Cowie, P.A., 1990, Determination of total strain from faulting using slip measurements, *Nature*, 838-840.

Sibson, R.H., J.M. Moore, and A.H. Rankin, Seismic pumping--A hydrothermal fluid transport mechanism, *J. Geol.Soc. London*, 131, 653-659, 1975.

Suess, E., Carson, B., Ritger, S., Moore, J. C., Jones, M., Kulm, L. D., and G. Cochrane, 1985, Biological communities at vent sites along the subduction zones off Oregon, in The Hydrothermal Vents of the Eastern Pacific: An overview, edited by M. L. Jones, *Bull. Biol. Soc. Wash., 6*, 475-484.

Turcotte, D., 1986, Fractals and fragmentation, *J. Geophys. Res., 91,* 1921-1926.

Vrolijk, P., Myers, G., and Moore, J. C., 1988, Warm fluid migration along tectonic melanges in the Kodiak accretionary complex, Alaska, *J. Geophys. Res., 93*, 10,313-10,324.

Vrolijk, P., 1987, Tectonically driven fluid flow in the Kodiak accretionary complex, Alaska, *Geology, 15*, 466-469.

Wintsch, R. P., Kvale, C. M., and H. J. Kisch, 1991, Open system, constant-volume development of slaty cleavage, and strain-induced replacement reactions in the Martinsburg Formation, Lehigh Gap, Pennsylvania, *Geol. Soc. Am. Bull., 103*, 916-927.

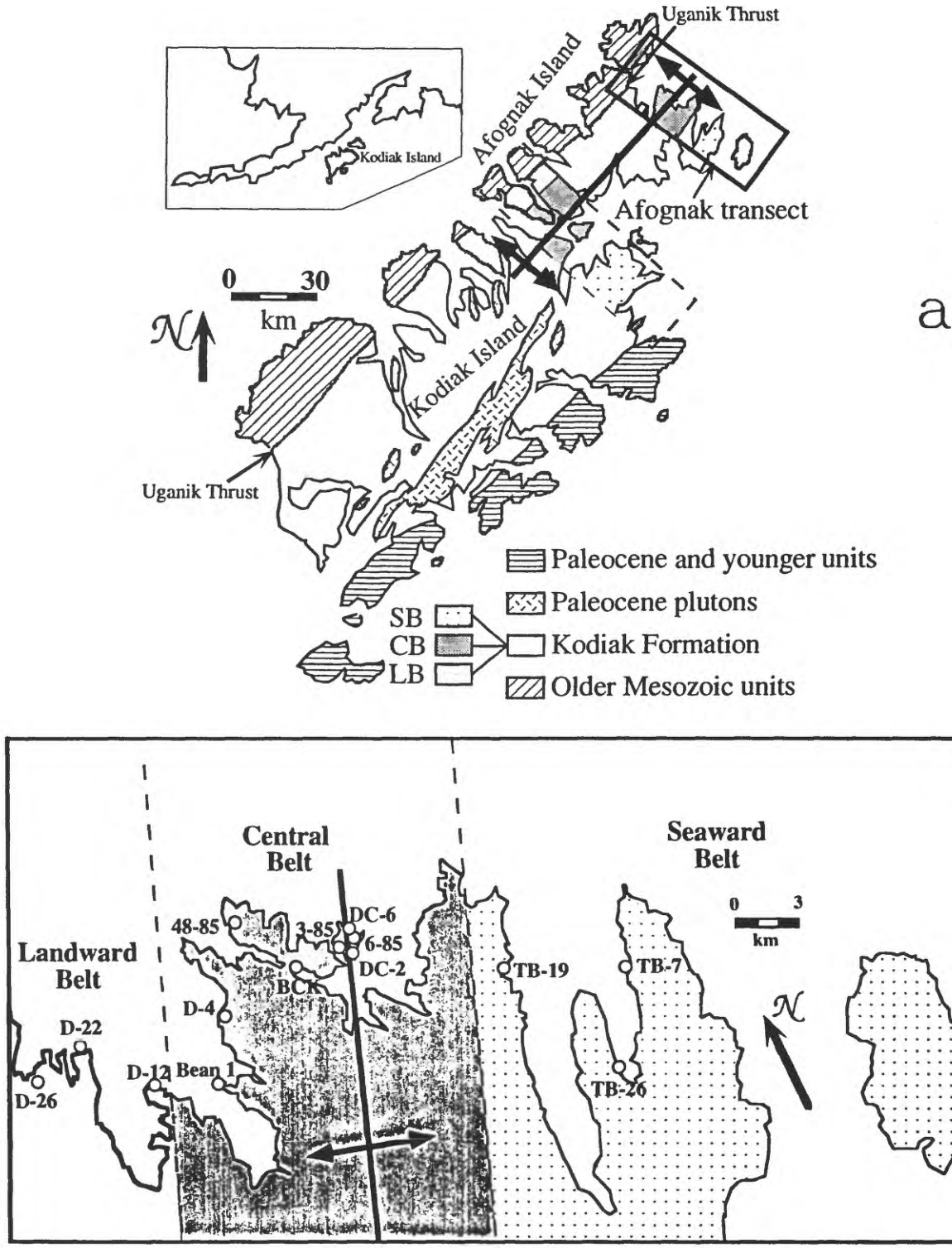



**Figure 1a.** Map of Afognak Island, Alaska (see insert), showing the Kodiak Formation and its division into the landward, central, and seaward belts. **b.** Map of the Kodiak Formation showing all sample locations.

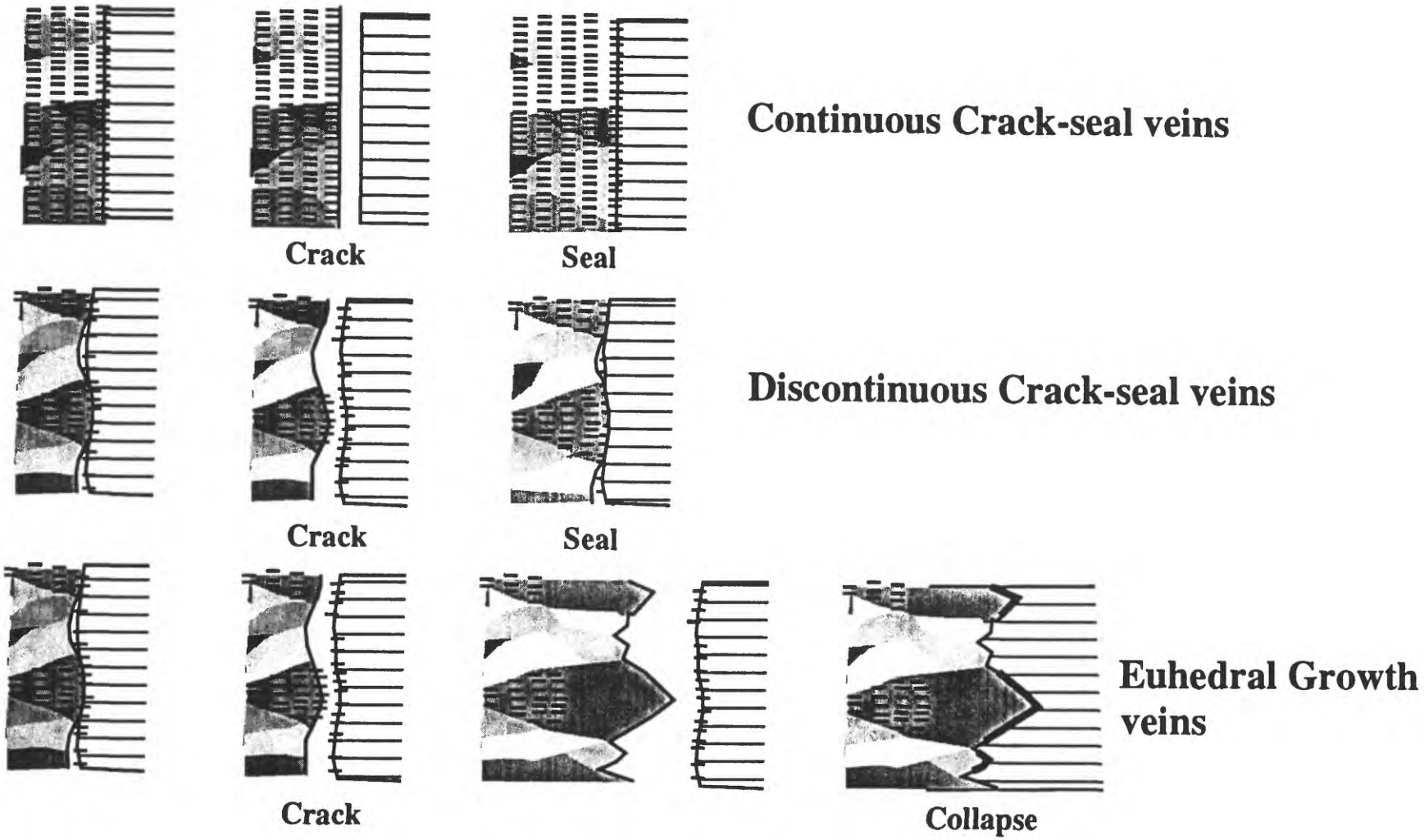





**Figure 2.** Schematic representation of quartz vein textures described in the text and in *Fisher and Brantley* [1992]. Continuously banded crack-seal veins contain bands of quartz separated by bands of chlorite inclusions. Quartz growth occurs on one side of the vein (quartz crystals shown as white to grey as they would appear under crossed polars) while chlorite (shown as tick marks) occurs from the opposite side from the wallrock. Each fracturing episode pulls chlorite from the wallrock. Discontinuously banded crack-seal veins only show chlorite bands along fast-growing quartz crystals which seal the fracture before a new fracturing event; slow-growing crystals never seal. Euhedral growth veins start as continuously or discontinuously banded crack-seal veins but develop into veins where quartz growth occurs predominantly into an open fluid-filled fracture. These veins show euhedral growth terminations which are paralleled in the matrix by insoluble residues (shown in black) where the terminations have jammed into the matrix after collapse.

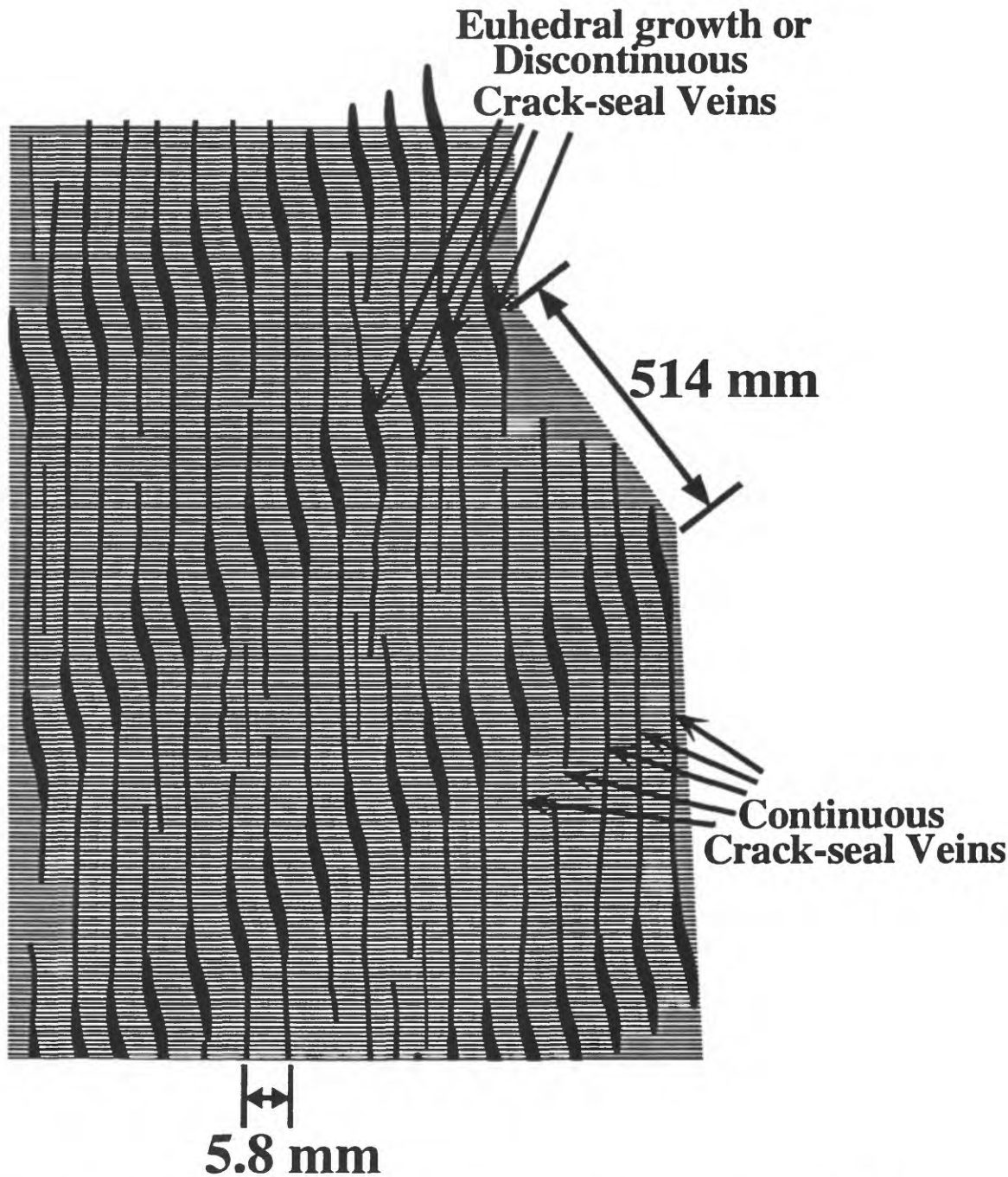



**Figure 3.** Schematic of the fracture network observed in the central belt of the Kodiak Formation. Distribution of vein thickness, vein spacing, and shear zone spacing is lognormal; median value for these parameters are indicated.

**Figure 4.** Photograph of outcrop showing distribution of thin veins and shear zones at location DC-2, Afognak Island.

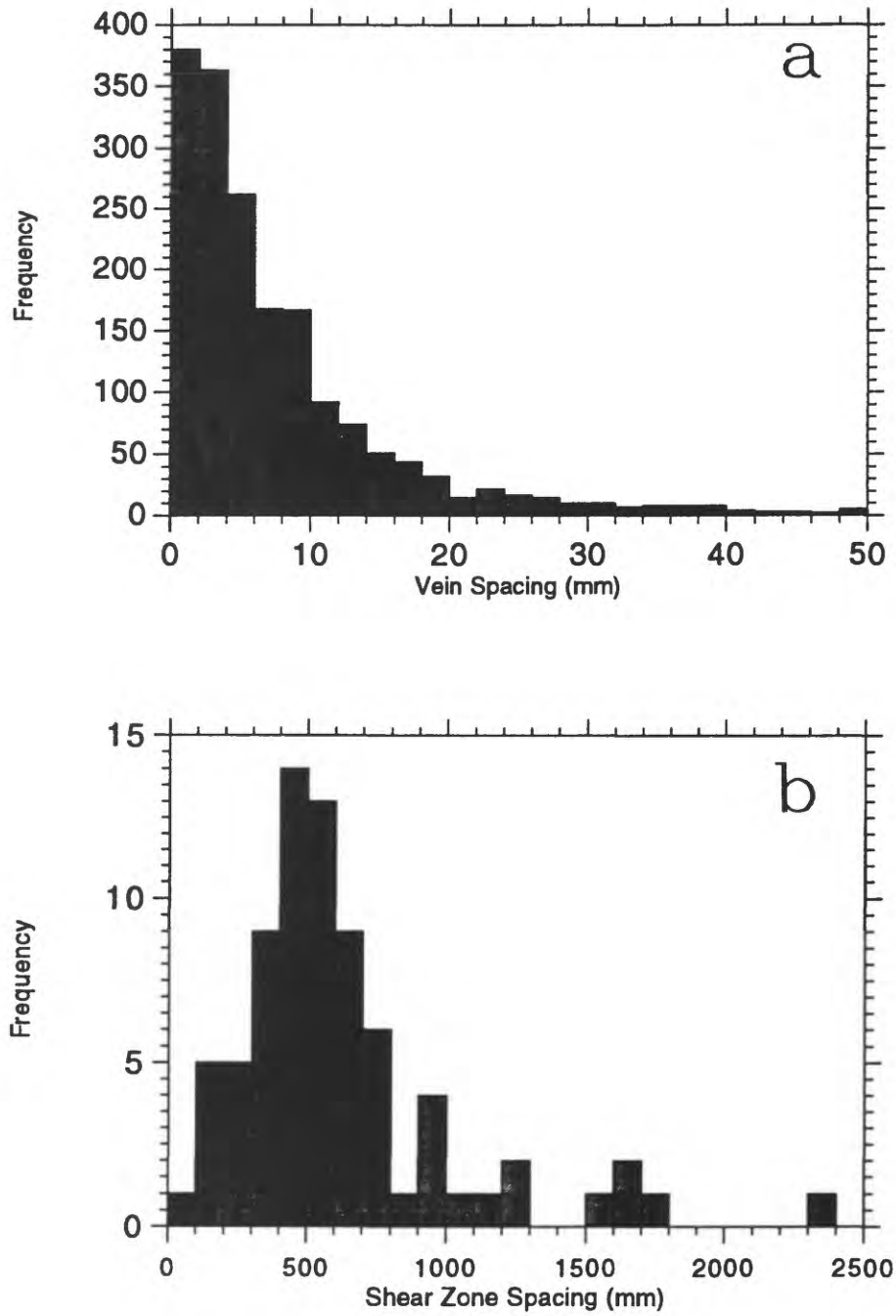



**Figure 5a.** Frequency histogram of vein spacing as measured on scanlines of $D_{2b}$ veins in slates throughout the central belt. Mean and median of the distribution are 12.22 and 5.81 mm respectively. **b.** Frequency histogram of spacing of shear zones throughout the central belt. Mean and median of the distribution are 622.8 and 513.9 mm respectively.

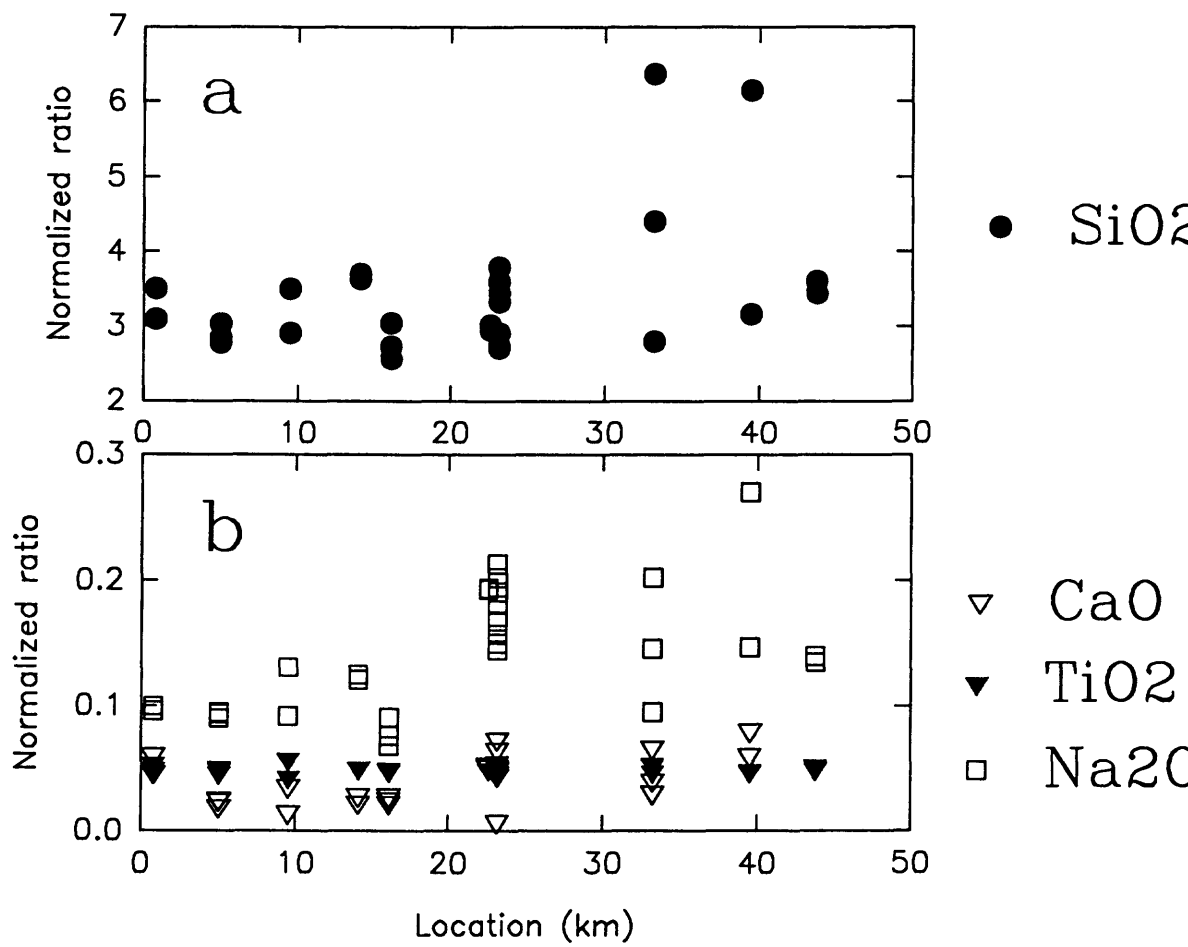



**Figure 6a.** Ratio of wt. % $SiO_2$ normalized with respect to wt. % $Al_2O_3$ throughout the landward (0-10 km), central (10-25 km), and seaward belts (25-45 km). Multiple data points indicate multiple samples at the same location. **b.** Ratio of wt. % CaO, $TiO_2$, and $Na_2O$ normalized with respect to wt. % $Al_2O_3$ throughout the transect. Samples which represent the deepest structural level lie between kilometers 15 and 25.

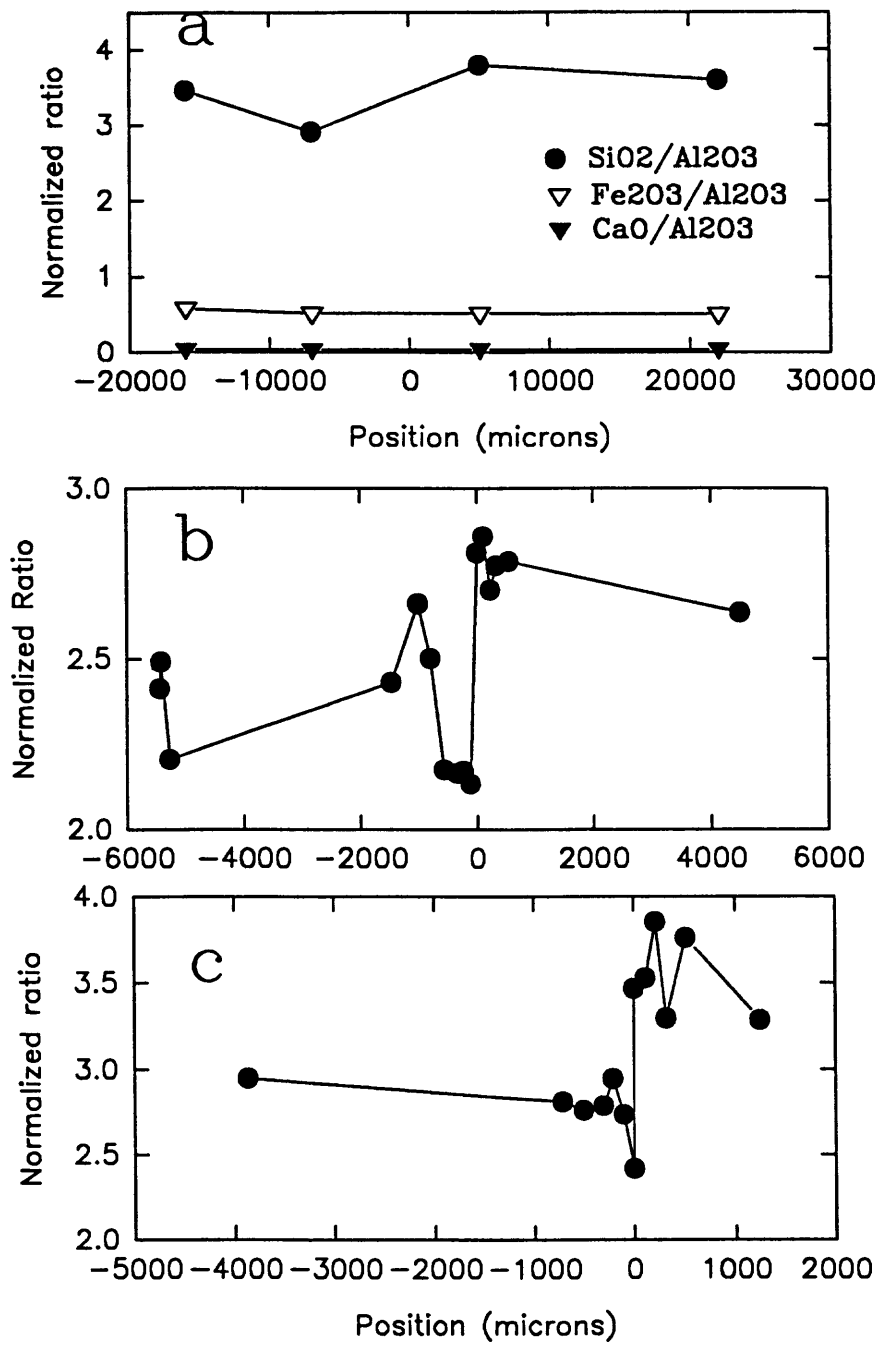



**Figure 7.** Ratio of wt. % $SiO_2$ (solid circles) normalized with respect to wt. % $Al_2O_3$ plotted as a function of position with respect to an euhedral growth vein in each of three samples, **a)** DC-2 **b)** DC-6 **c)** BCK. Wt. % $Fe_2O_3$ and CaO normalized with respect to wt. % $Al_2O_3$ is also plotted. Note that scales differ for each vein. For every plot, the euhedral growth quartz vein was located at position 0. For every case, the side of the vein with euhedral terminations is on the left (negative position). Measurements plotted in (a) were determined with whole rock analysis; measurements plotted in (b) and (c) were determined from 110 μ x 90 μ rastered analyses under electron microprobe.

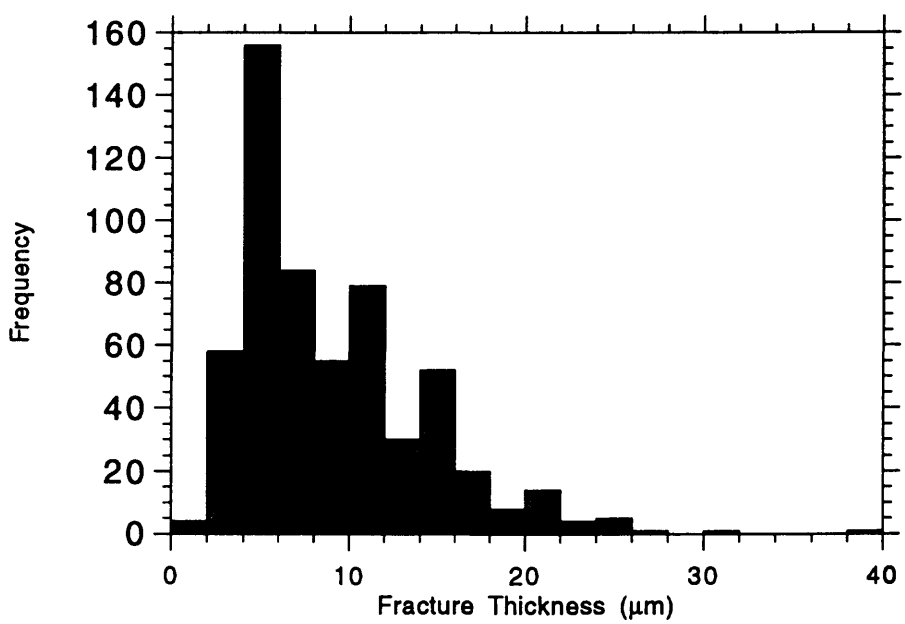



**Figure 8.** Frequency histogram for thickness of quartz bands in continuously banded crack-seal veins for samples observed in five localities. Mean and median of the distribution are 9.0 and 7.1 μm respectively.

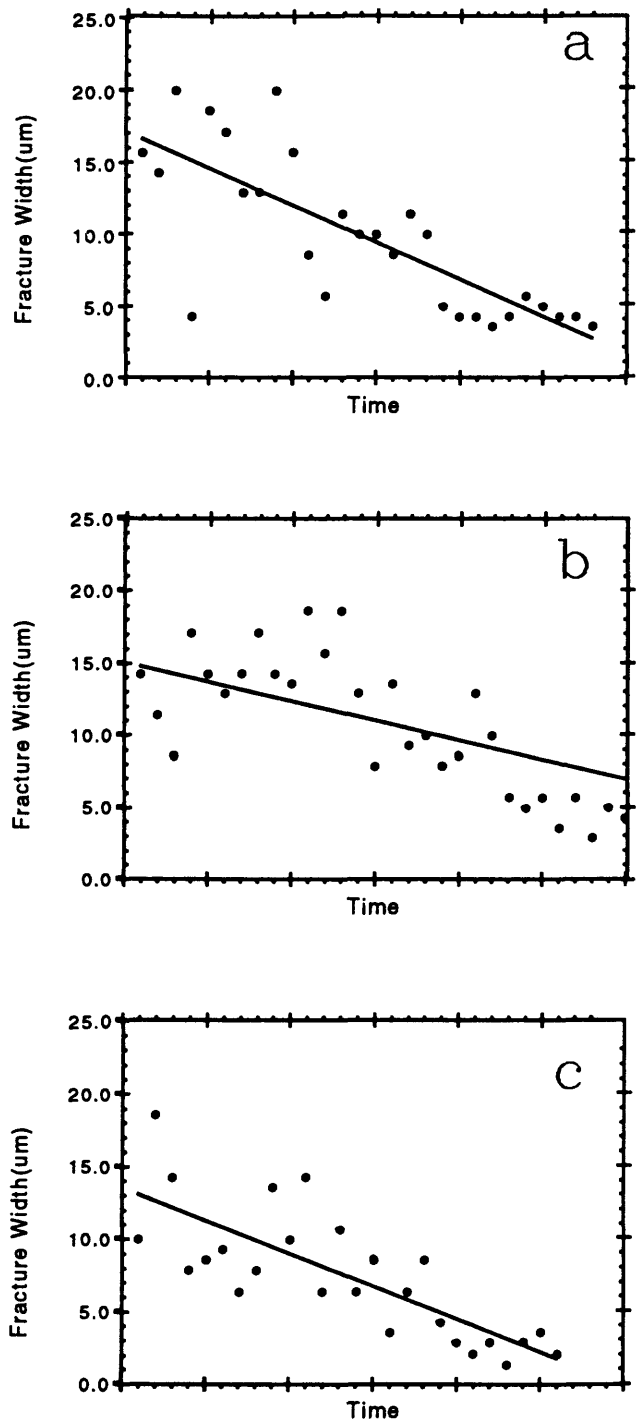



**Figure 9.** Thickness (μm) of quartz bands plotted from the first band grown (measured on the side of the vein where quartz grains are smallest) to the last band grown (measured on the side of the vein where quartz grains are largest). Plots in (a), (b), and (c) represent measurements on three segments of the same crack-seal vein and indicate the extent of variability within one vein. The mean and standard deviation in each case is: a) 8.20 ± 1.80 μm, b) 6.91 ± 1.81 μm, c) 6.12 ± 1.96 μm. Lines represent a linear regression through all data.

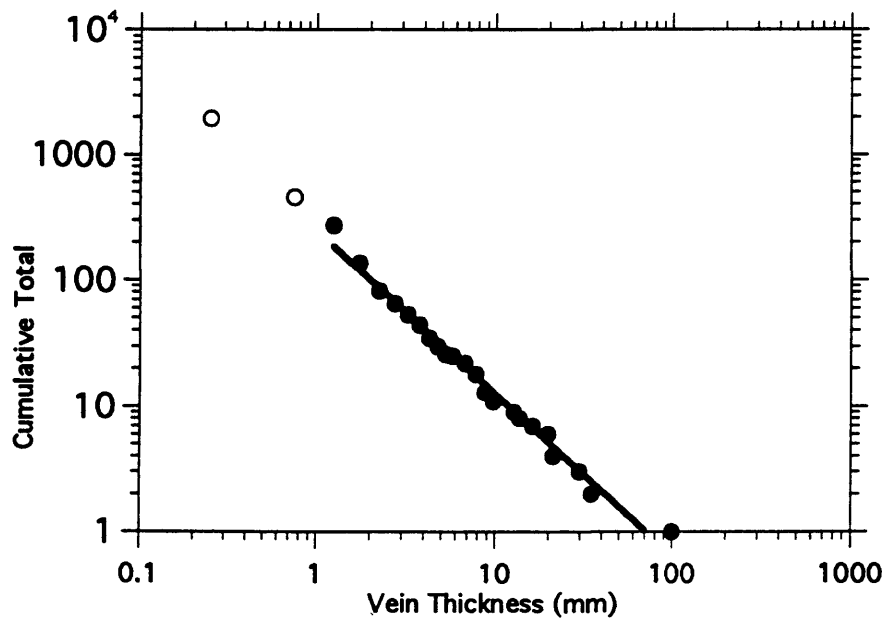



**Figure 10.** Cumulative total (N(X)) of veins exhibiting thickness >X (mm). The slope of the regression line is 1.3. Solid data points were used in the linear regression. Note that the data point for the smallest value of X represents the mean of all vein thickness values as measured under the microscope for veins < 1 mm (see text). All other data points represent measurements made in the field and then corrected for vein orientation.

# The Nature of High Fluid Pressure in Sedimentary Basins

**Terry Engelder, Suzanne D. Weedman, and John T. Leftwich, Jr.**
Department of Geosciences, The Pennsylvania State University
University Park, Pennsylvania 16802

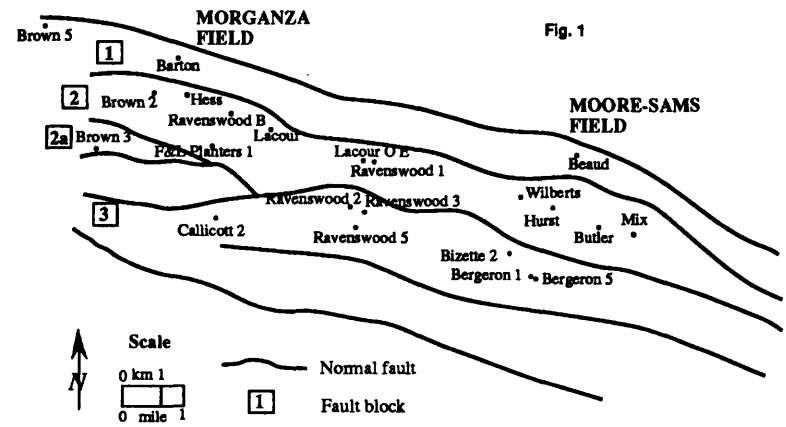
Any study of the mechanical involvement of fluids in faulting will undoubtedly consider the effect of high fluid pressure. While measurements of fluid pressure about faults are rare, some of the best data come from oil and natural gas fields in sedimentary basins (e.g., the Gulf of Mexico). Such detailed pressure data give an insight into the distribution of abnormal pressure in and around fault zones where high fluid pressures are quasi-static relative to the earthquake cycle. For quasi-static fluid pressures, the nature of the pressure seal is an issue. The purpose of this paper is to describe the change in fluid pressure when abnormal pressure is encountered in sedimentary basins. In particular, focus is on the question of whether or not high fluid pressures are any different along fault zones than within intact rock at some distance from the fault zones.

Geopressure profiles were constructed largely from original shut in bottom hole pressure (BHP) measurements recorded in numerous wells within various oil and gas fields of a sedimentary basin. Such pressure data were recorded by the operators during the completion and production phases of each well and are available through a Houston based company, Petroleum Information Service (PI). In some cases repeat formation tester (RFT) and other wireline (WLT) pressure data supplemented the BHP data.

Early compilations of fluid pressure data from the Gulf of Mexico suggest that abnormal pressure increases in a rather nonsystematic manner at a depth where shales become more common in the stratigraphic column. In the literature (e.g., Magara, 1978; Chapman, 1980), a common model for pressure increase in Tertiary rocks is well expressed by a figure from Wallace et al (1979) which shows that below the hydrostatically pressured zone, pore pressure increases with depth so that the geopressure profile is a concave curve that grades into a lower convex curve. The result is a geopressure profile with the shape of a "lazy S". Such geopressure profiles are characterized by lack of sharp boundaries, in contrast to geopressure profiles drawn through seals bounding pressure compartments (e.g., Hunt, 1990; Weedman et al., 1992). In fact, Chapman (1980) cites the "lazy S" shape of the geopressure profiles as evidence that there are no seals (i.e., distinct low permeability horizons relative to rocks above and below) at the top of overpressure in thick shale sections of the Gulf Coast.

Construction of a geopressure profile assumes that high quality data from each well in individual fault blocks follow the same general trend and that differences in original pressures due to rollover anticlines were relatively small. Special care was used to select original pressures and exclude low pressures taken after a particular zone had experienced some production drawdown which is indicated by data falling off of a geopressure profile by more than 2 MPa. Pressure points that were determined to be faulty or drawn down were left off of the final plots. Detailed pressure records were not available for any of the BHP data, therefore evaluating the quality of each test was not possible.

Tuscaloosa Trend The first example of abnormal pressure about fault zones is taken from the Tuscaloosa trend in Louisiana (Weedman et al., 1992). The very porous sandstones of the lower Tuscaloosa Formation (Upper Cretaceous) produce gas in the depth range of 5400 to 6400




meters south of the Lower Cretaceous shelf margin (Fig.1). In the **Moore-Sams** and **Morganza** fields, the formation is cut by several east-west trending, southward dipping growth faults (Figure 1), and comprises shales and both fining- and coarsening-upward sand bodies interpreted as distributary channels, distributary-mouth bars, and offshore bars. A transition to overpressure occurs within the lower Tuscaloosa Formation in these two fields. Reservoirs within the lower Tuscaloosa Formation < 10 km to the north are normally pressured while reservoirs in the same formation < 10 km to the south are entirely overpressured (McCulloh and Purcell , 1983).

Moore-Sams Field All pore pressures measured in fault block 1 fall on a hydrostatic gradient to the ground surface; the abnormal pressures in blocks 2 and 3 of this field follow a local hydrostatic gradient (Figure 2). The overpressures of blocks 2 and 3 follow a local hydrostatic gradient down to depths of 5850 m reaching magnitudes of 82.6 MPa. The top of the transition zone from normal to overpressure occurs in the depth range of 5620 m to 5683 m, while the top of the lower Tuscaloosa Formation is displaced down by 105 meters from block 2 to 3. Therefore, the variation in the depth to the top of overpressure between fault blocks 2 and 3 is considerably less than the displacement of the strata across the fault.

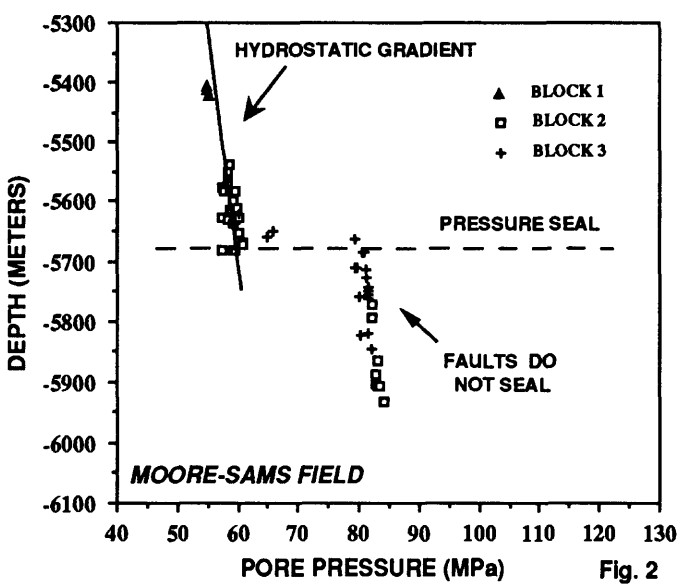


Fig. 2

It is important to note that the magnitude and gradient of abnormal pressures in blocks 2 and 3 are very similar, even though stratigraphy differs across the fault as indicated by the displacement of both the top of the lower Tuscaloosa Formation (by ~105 m) and the shale break at the base of the formation (by ~180 m ), as well as in the basinward thickening (~ 78 m) of the formation. The fault that separates the two blocks, therefore, does not act as a pressure seal, because overpressures are nearly equal on both sides of the fault at the same depths. There is apparent fluid communication across the fault and within the overpressured sandstones of this field. Communication through the fault zone indicates that the effective stress within the fault zone is the same as found within intact rock on either side of the fault.

Morganza Field Overpressures in the Morganza Field wells reach magnitudes as great as 117 MPa at depths of 5975 meters (Figure 3). The pore pressures through overpressured sand-

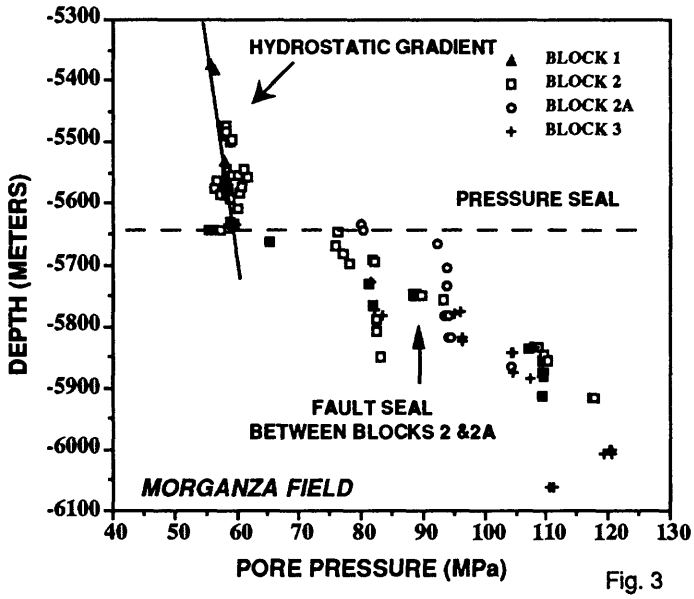



Fig. 3

stone-dominated intervals in the Morganza Field generally follow local hydrostatic gradients, but jump as much as 26 MPa across certain shaley intervals. The growth faults in this field divide up the reservoir into at least five blocks; four of them have wells with RFT pressure data. As in the Moore-Sams Field, all pore pressure data from fault block 1 are normal, i.e., they lie on a hydrostatic gradient to the ground surface, while the overpressures encountered in fault blocks 2 and 3 are of a much greater magnitude than in the Moore-Sams Field. The pore pressures in the OE Lacour well and in part of Brown 3 well reach magnitudes of 83 MPa and 94 MPa, respectively, at depths of about 5800 meters, while pore pressures in the Brown 2 well follow a series of stair-step local hydrostatic gradients, reaching magnitudes as great as 117 MPa at depths of 5975 meters.

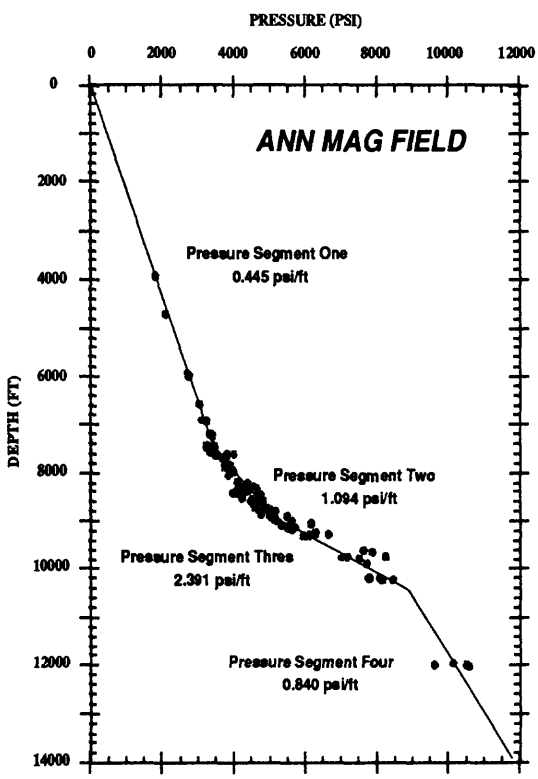



Fig. 4

The pressure data from the Morganza Field are more complex than the data from the Moore-Sams; a great deal of variation in pore pressure exists in the depth range of 5675 to 5850 m, suggesting that several pressure seals may exist, not only along certain faults (between blocks 2 and 2a), but even within fault blocks (i.e., 2 and 3). Here faults serve as pressure seals which are zones of sufficiently low permeability to maintain a measurable vertical or lateral pressure anomaly. Effective normal stress within such sealed fault zones can be sizable.

South Texas Geopressure profiles for many fields in South Texas show several segments having linear trends. As a consequence, we computed a best fit or we hand-fitted a line to linear segments of the geopressure profiles.

A geopressure profile constructed for the **Ann Mag** field allows us

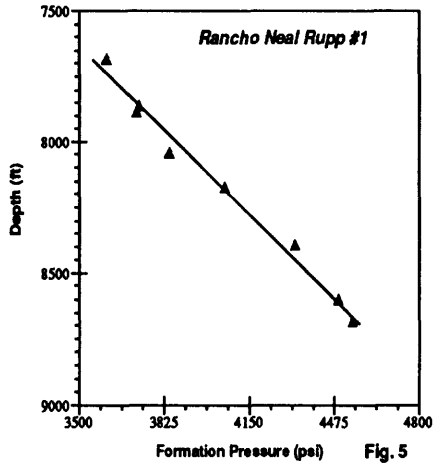


Fig. 5

to identify the top of abnormal pressure at a depth of approximately 2286 m. The most striking feature of the Ann Mag pressure-depth profile is that it is divided into four linear segments with characteristic pressure gradients (Figure 4). A normal pressure gradient of 0.445 psi/ft (10.1 MPa/km) is found down to the top of abnormal pressure. Immediately below the top of abnormal pressure is a normally compacted section with a pressure gradient of 1.094 psi/ft (24.7 MPa/km). Below the top of undercompaction is a section with a high gradient of 2.391 psi/ft (54.1 MPa/km) down to a depth of approximately 10500 ft (3200 m). The deepest section has a pressure gradient of about 0.84 psi/ft (18.9 MPa/km) which is poorly constrained because of the small number of data points. When encountering these linear pressure-depth segments in other fields, they are labeled ONE through FOUR moving downward in a field. The mean top of the zone of undercompaction in the Ann Mag field at a depth of 8980 ft (2737 m) is 1480 ft (451 m) below the mean top of abnormal pressure in this field.

The skeptic will undoubtedly argue that the raw data in figure 4 follows a curve more like the "lazy S" of Chapman (1980). However, figure 4 is a compilation over an entire field where fault zones do not act as seals. Pressure data from single wells serve as the best evidence for linear segments in pressure-depth profiles as opposed to nonlinear pressure increases. Pressure segment TWO for the Rupp 1 well is particularly well defined (Figure 5). A computer calculated curve fit to the data points comprising this segment indicates a linear curve fit with a regression coefficient of 0.993. The abrupt change in pressure gradient that occurs in going from the normally pressured section to abnormally pressured section is also consistent with linear pressure-depth segments. Observations also indicate that the change from normal compaction to undercompaction occurs abruptly, rather than gradually and that this abrupt change is accompanied by an abrupt change in abnormal pressure gradient as seen in going from pressure segment TWO to pressure segment THREE in the Ann Mag field (Leftwich and Engelder, 1993).

In the **Alazan** field the top of undercompaction occurs at essentially the same depth as the top of abnormal pressure. Three segments exist with the middle segment having a pressure gradient as high as that of segment THREE in the Ann Mag field. Segments in the Alazan field are equivalent to segments ONE, THREE, and FOUR in the Ann Mag and field. Segment TWO is missing because the top of abnormal pressure and the top of undercompaction are coincident.

In summary, the Ann Mag field shows four linear pressure segments with the top of undercompaction defining the boundary between segments TWO and THREE. Although segment FOUR is poorly constrained in the Ann Mag field due to lack of data, a linear segment FOUR is well constrained in other Tertiary fields. A field with four pressure segments is called the Ann Mag-type field. In contrast the Alazan field had three pressure segments so, fields with three pressure segments are called Alazan-type fields (Leftwich and Engelder, 1993).

Both types of Tertiary geopressure profiles are distinct from those encountered in older rocks of the Gulf of Mexico where a pressure seal cuts through thick Cretaceous sand bodies within the deep lower Tuscaloosa trend of Louisiana (Weedman et al., 1992). The distinction in geopressure profiles is attributed to differences in lithology, stratigraphy, and seals. Geopressure

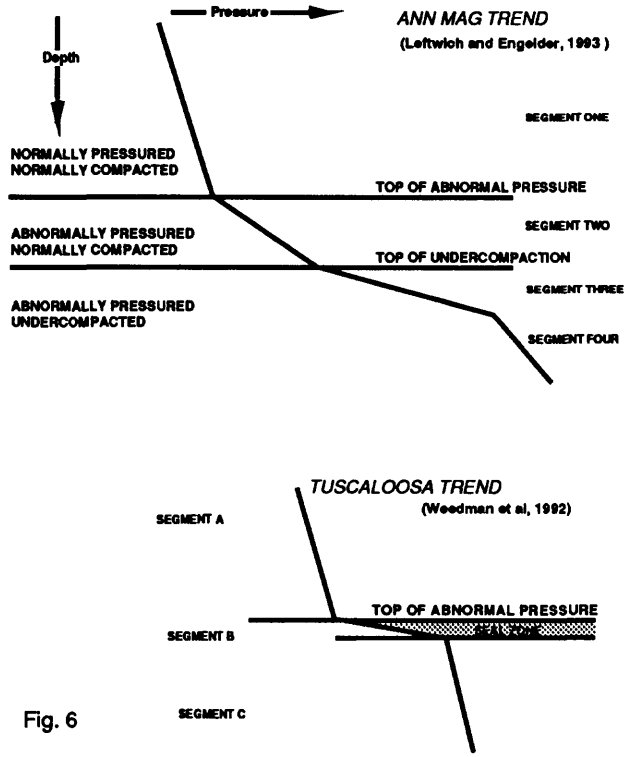



Fig. 6

profiles in the deep Tuscaloosa sandstones are characterized by a much thinner pressure transition (~ 20 m) (i.e., a pressure seal a la Powley, 1990) where the local pressure gradient is as much as 0.43 MPa/m (pressure segment B of Figure 6). We know less about the shallower Tertiary geopressure profile above the Tuscaloosa trend, but, we know that it is overpressured at depths of ~4,200 m, with a regression to normal pressure at the top of the lower Tuscaloosa Formation. Tuscaloosa-type geopressure profiles are characteristic of pressure compartments (Hunt, 1990) with pressure segments A and C having a gradient of about 0.01 MPa/m which is characteristic of freely communicating pore fluid in relatively permeable rocks. Above pressure segment A of the Tuscaloosa-type geopressure zone, there is a shallower abnormal pressure zone which could be characterized by either an Ann-Mag or Alazan geopressure profile. The present data are too sparse to distinguish between the two.

In comparing Tuscaloosa-type and the two Tertiary geopressure profiles, pressure segments A and ONE are the same. While pressure segment C has a hydrostatic gradient, it is distinguished from pressure segment A by abnormal fluid pressures and it does not have the same pressure gradient as pressure segment FOUR. Likewise, pressure segments B, TWO, and THREE are all distinct from each other (Figure 6). Thus, we have recognized that geopressures within the Gulf of Mexico are characterized by at least six different pressure segments. The following table gives ranges for pressure gradients for each of the six pressure segments in terms of kPa/m:

| Segment: | ONE & A | TWO | THREE | FOUR | B | C |
|---|---|---|---|---|---|---|
| Gradient : | 10.1-10.2 | 20 - 34 | 45 - 100 | 19 - 21 | 450 + | 10.1-10.2 |

In conclusion, one interpretation of the pressure data is that in non-sealing fault zones effective stress is about the same as that in the country rock and, hence, such fault are readily able to slip. In contrast, in sealing fault zones, the effective stress can be considerably higher than in flanking country rock and, hence, these fault more likely to remain locked.

References

Chapman, R.E., 1980, Mechanical versus thermal cause of abnormally high pore pressures in shales: American Association of Petroleum Geologists: v. 64, p. 2179-2183.

Hunt, J.M., 1990, The generation and migration of petroleum from abnormally pressured fluid compartments: American Association of Petroleum Geologists Bulletin, 74, p. 1-12.

Leftwich, J.T., and Engelder, T., 1993, The relationship between undercompacted shale and abnormal pore pressure in the Gulf of Mexico Basin: American Association of Petroleum Geologists Bulletin, (submitted).

Magara, K., 1978, Compaction and Fluid Migration — Practical Petroleum Geology: New York, Elsevier Publishing Company, 319 p.

McCulloh, R. P. and Purcell, M. D., 1983, Hydropressure tongues within regionally geopressured lower Tuscaloosa Sandstone, Tuscaloosa trend, Louisiana: Trans. Gulf Coast Assoc. Geol. Soc. 33, p. 153-159.

Wallace, R.H., Jr., Kraemer, T.F., Taylor, R.E., and Wesselman, J.B., 1979, Assessment of geopressured-geothermal resources in the northern Gulf of Mexico basin: U.S. Geological Survey, Circular 790, p. 132-155.

Weedman, S. D., Guber, A. L., and Engelder, T., 1992, Pore pressure variation within the Tuscaloosa trend: Morganza and Moore-Sams Fields, Louisiana Gulf Coast: Journal of Geophysical Research, v. 97. p. 7193-7202.

# SURFACE GEOPHYSICAL METHODS TO DETECT FLUIDS RELATED TO FAULTING

Donna Eberhart-Phillips[1], William D. Stanley[2], and Brian D. Rodriguez[2]
[1]U.S. Geological Survey, 525 S. Wilson Ave., Pasadena, CA 91106
[2]U.S. Geological Survey, Box 25046, MS964, Denver Federal Center, Denver, CO 80225

## INTRODUCTION

Faults generally juxtapose blocks of differing material properties that can separately be imaged with surface geophysical methods. But if we consider the fault zone itself, the problem becomes more difficult. The fault zone may be less than 1 km wide, highly fractured and filled with fault gouge and fluids. The combination of separate fractures, breccia, clay and particles in the gouge, and presence of fluids at variable pore pressures provide large contrasts in physical properties such as P and S-wave acoustic velocities, electrical resistivity, density, and magnetic susceptibility. However, the narrow aspect of this physically anomalous zone makes sensing with surface geophysical methods difficult, particularly when the seismogenic zone of most interest is greater than 5 km deep. Subsurface measurements using boreholes that straddle the fault are generally more successful, but in many regions of important faults such borehole pairs are not available. In addition to the fluids that occur in faults and control their stress behavior, fluids play a key role in where and when faults develop. Geophysical methods can be used to map fluids within existing fault zones or the occurrence of regional fluid environments that control the development of faults. We will review a variety of methods and some case histories involving geophysical methods applied to measuring the specific case of fluids in faults and in reservoirs that control fault development. Also, we provide some model simulations to suggest techniques to be used and the degree of difficulty involved in investigating faults and the fluids associated with them.

## GEOPHYSICAL TECHNIQUES

Material properties that can be imaged from the surface (or from boreholes) that are affected by static fluids are compressional seismic velocity ($V_p$), shear seismic velocity ($V_s$) or $V_p/V_s$ ratio (related to Poisson's ratio), seismic attenuation, and electrical resistivity. In addition, fluid in motion within a fault zone or a confined reservoir in contact with a fault generates streaming electrical potentials that can be measured at the surface or from boreholes.

### Seismic Methods

Seismic imaging can be done using natural local sources (earthquakes) or active sources (blasts or impulsive/periodic impactors such as Vibroseis). The seismic energy from active sources can be recorded using refraction or reflection methods, and also used for detailed ray path tomography across faults. Earthquake energy from local aftershock sequences can be used for detailed inversions of 3-D structure and hypocenters. This is commonly referred to as local earthquake tomography. The location and origin time of local earthquakes are unknown parameters and must be included in the inversion because the earthquakes are within the volume being imaged. This is particularly true when one is trying to image narrow details and not just broad velocity variations. Each arrival time will provide velocity information along the particular ray path from that earthquake to that station. Ray paths bend toward higher velocity material and so will become more geometrically complicated (i.e.: not simple, planar straight or curved paths) where there is a strong velocity gradient. Thus, in order to best image a narrow low-velocity zone, it is important to use a fairly accurate 3-D ray tracer and to redo the ray tracing for every iteration of the velocity inversion.

Active source seismic imaging is usually done along profiles containing the receivers and the sources, so 2-D models are usually constructed. Standard, exploration style reflection

surveys use source-receiver separations of a few tens of meters and dynamite shots or mechanical impactors. Each source-receiver distance is small so that only near-vertical ray paths are utilized. For this reason "vertical-incidence" surveys like those used in petroleum exploration are most useful for imaging subsurface structures with small dips, such as might be the case for a fluid reservoir in contact with a fault, but not for the fault itself (Mooney and Brocher, 1987). Refraction seismic methods involve use of long offsets between shots and receiver, depending upon the problem. Refraction surveys involve ray paths that have refracted from interfaces, so that the method usually works best with nearly horizontal geological units. Calculation of delay times under the assumption that the geology is pseudo-horizontally layered can allow mapping lateral variations in velocity that can be used to locate low-velocity zone near faults due to breccia, gouge zones and fluids. A variation of refraction recording known as wide-angle recording uses a wide range of source receiver offsets so that near-critical reflections with high energy can be recorded from important horizons. Velocity models can be obtained from refraction or wide-angle records by inversion or forward modeling using ray-tracing or finite-difference algorithms. The location of a fault can be inferred from the velocity contrast where a fault separates differing lithologic blocks, but the wide-angle technique is less applicable to direct studies of the narrow fault zone itself. However, wide-angle experiments may be important for study of mid-crustal fluid regions that control the development of regional detachments and faults.

While travel-time methods are most commonly used to interpret seismic velocity, there are other seismic techniques which look at the waveform details. Of particular interest for fault zones is the analysis of trapped waves (Ben-Zion, et al., 1992). Focusing of seismic energy in fault zones has been modeled with synthetic seismograms by Cormier and Spudich (1984). Differential attenuation due to physical property variations affect the waveforms. Zucca, et al. (1993) have developed inversions for attenuation using the pulse-width of the initial pulse. Hough et al. (1991) and Gou et al. (1992) have developed a method to infer attenuation that uses the spectra of the first few seconds of each earthquake seismogram.

Shear-wave velocity anomalies in the area of faults have been analyzed in terms of dilatancy models, but little direct use of shear-wave recording using refraction or wide-angle seismic geometries has been done near faults to directly study fluids. In the New Madrid seismic zone, detailed refraction-wide-angle reflection surveys have been done to investigate the "Blytheville arch", a linear zone possibly connected with basement faults that corresponds to seismicity trends (W. D. Mooney, writ. comm., 1993). This zone is one of high attenuation (Hamilton and McKeown, 1988) and geological evidence suggests high-pore-pressures have been active in the past and recent history (Diehl and McKeown, 1991). The survey indicated a pronounced low-velocity zone, but the region occurs at depths greater than 5 km beneath a complicated sedimentary sequence. Hence the probability of deducing fluids states is low using P-wave arrivals only and S-wave arrivals have not yet been analyzed. The difficulty in studying a narrow fault zone with high shear-wave attenuation is significant, and it may be proven in the future that properly designed surveys that use energy which refracts through the fault zone at optimal angles might be valuable in studying the fluid content of the fault.

Electrical Methods

Electrical methods are valuable for imaging fault zones because many of the properties of the faults make them good conductors. This is the case when fault gouge has large percentages of clay, which contains bound but electrically active fluids, as well as free fluids. Also, some fault zones contain conductive minerals such as sulfides and graphite which make them extremely good conductors. Fault zones can be more resistive than their surroundings when they are filled with precipitated minerals such as calcite or quartz that have high resistivity. These resistive fault zones can be considered potentially "healed", whereas conductive zones can be considered potentially "active". Faults within crystalline rocks generally have resistivities orders of magnitude lower than the surrounding rocks. However, faults that extend

upward into conductive sedimentary rocks are more difficult to image because the sedimentary rocks also contain large amounts of clay and fluids, similar to those within the fault.

In general it must be assumed that it would be impossible to distinguish actual volumetric fluids within shallow parts of the fault from extensive clay-rich gouge using most electrical methods. However, fault gouge usually has a lower limit in resistivity of approximately 3 to 10 ohm-m in the presence of normal meteoric water (possessing few ions, making it high-resistivity). Heavily mineralized fluids within the fault zone can significantly reduce resistivities in the fault below this range. Mineralized fluids rich in Na, Ca, Cl, $CO_3$ are common in active fault zones. If upwardly moving, mineralized fluids are invading the fault, electrical imaging methods can be used to map the distribution of the fluids along a fault and within reservoirs lateral to the fault. As we discuss below, the self-potential method may be used to directly investigate fluids if they are in motion.

For deeper regions of the fault (below about 5-8 km) fluids that are associated with nearby permeable rocks are restricted due to decreased porosity, and prograde metamorphism may be the source of most of the fluids. Because of the larger depth-to-width aspect of deeper parts of faults the only possible way of sensing these fluids with electrical geophysical means is if they provide contact of the surface with a conductive deep crustal or mantle regime. This fluid-related, short-circuit can drastically affect certain low-frequency magnetotelluric (MT) data, as discussed below.

A variety of methods have been used to obtain electrical images of fault zones and regional features that may be important to fault generation. As in the seismic methods, both artificial and natural energy sources can be utilized. For detection of shallow features of fault zones (including the fluid content), direct current (DC) methods have been used. The DC methods involve injection of current into the earth and measurement of electric fields with a specific electrode geometry. Inductive methods, known as electromagnetic (EM) techniques, can also be used to measure resistivities across fault zones and obtain subsurface images. These methods use an artificial, harmonic or time-domain source to induce currents in the fault zone and surrounding medium that are measured in terms of either magnetic or electric (or both) components. One advantage of EM methods is that they are highly sensitive to conductive zones, especially to long, vertical zones with the appropriate geometry of transmitter and receiver. Natural source EM methods include the magnetotelluric and geomagnetic variation (GMV) methods.

The MT method measures the frequency variations in electrical and magnetic fields produced by ionospheric disturbances and lightning, generally over 4-7 decades of frequency from 0.001 Hz to 10,000 Hz. This large frequency range, combined with the large dynamic range of resistivity of earth materials, has made the MT method very popular in tectonic studies involving fault zones and orogenic belts. A controlled, artificial source is sometimes used in the 20-20,000 Hz frequency (audio) range and is called the controlled source audio-magnetotelluric method (CSAMT). When a well-located transmitter is utilized the CSAMT method can provide excellent lateral and vertical resolution of the relatively shallow parts of a fault zone.

The GMV method employs an array of magnetometers to study current concentrations caused by variations in earth resistivity. Although not often utilized, this technique can be used to locate conductive faults when adequate station spacing and recording frequencies are used. Such an array can also be used to study magnetic signals produced by fluids flowing through the porous zones in a fault, although this is more easily done in most environments by monitoring electrical fields. There have been a number of reported successes in measuring transient electric fields associated with active faults, generally attributed to movement of fluids in and around the fault. Corwin and Morrison (1977) recorded self-potential (SP) anomalies associated with an earthquake south of Hollister, CA and attributed the anomalies to transient fluid flow associated with pre-quake stresses. Similar observations and interpretations have been made in the Soviet Union and China (Sobolev et al., 1975; Fung-Ming, 1976). Murakami et al. (1984) have extensive evidence for SP anomalies on earthquake faults in Japan and also

attribute the anomalies to motion of fluids in the fault and nearby rocks. Thus, it is probable that SP surveys represent one of the only known methods to study fluids in motion within fault zones. Fitterman (1979) has provided theory for calculation of SP near vertical contacts.

## EXAMPLES

It can be assumed that although many faults will appear as low-velocity, low-resistivity features in higher-velocity, higher-resistivity surroundings, many important faults will not have such contrasts. However, it is likely that the correlation of low-velocity and low-resistivity will hold where there is significant fault gouge. For healed faults that are cemented with calcite or quartz, the faults may appear transparent to both seismic and electrical methods, but this is diagnostic information for the role of fluids in faults as well.

One example of a possible transparent fault section is the Anza area of the San Jacinto Fault which is covered by an excellent local seismic array. Scott (1992) used the Anza data in a detailed 3-D velocity inversion and did not find an obvious low-velocity signature. This probably means that there is a limited width gouge zone associated with this fault or that this section is temporarily healed. Studies of the Parkfield segment of the San Andreas Fault (SAF) indicates variations in physical properties that are dominated by the contrast of Salinian granitic rocks adjoining Franciscan/Great Valley sedimentary rocks (Eberhart-Phillips and Michael, 1993), however there is a suggestion of a low-velocity zone (LVZ) along the fault, particularly at shallow depths (Michelini and McEvilly, 1991). As summarized by Mooney and Ginzburg (1986), substantial LVZs have been found on the SAF south of Hollister and along the Calaveras fault. These are 1-3 km wide features inferred to extend to at least 6 km and possibly throughout the crust. Narrower fault zones within these features are difficult to delineate although Lin and Roecker (1993a,b), using more accurate ray tracing image a narrower LVZ within the wide LVZ.

MT studies completed by Stanley (unpublished data, 1993) show that large electrical contrasts occur at Parkfield between the two rock types juxtaposed across the fault, but on some profiles there is indication of a conductive zone related directly to the fault extending to depths of up to 10 km. Park et al. (1991) have interpreted regional, low frequency soundings across the Parkfield segment of the SAF to require a conductive zone that extends through the crust. Such a feature requires the presence of fluids at mid- and lower-crustal depths. Mackie et al. (1993) also require such a through-crust conductive zone near Loma Prieta using sparse MT soundings and regional data, although Rodriguez et al. (1993) have not been able to substantiate this feature with more detailed data.

The question of a mid- or lower-crustal fluid zone is very important to behavior on the San Andreas and other deeply penetrating fault zones. In a typical active fault in crystalline or metamorphic rocks a clay rich gouge of various zeolite facies minerals is very conductive (<10 ohm-m) and low-velocity (< 2 km/s P velocity). The rocks in an active fault zone may have undergone varying degrees of cataclasis, producing tectonic breccias in instances of low normal stress with higher stresses producing higher proportions of crushed material and in the highest stress areas forming mylonites (Miyashiro, 1973). Short of the recrystallization involved in mylonite formation the wide crushed zone may have distributed clay minerals due to active hydrous and thermal alteration of the broken minerals. The currently active fault trace may be quite narrow and contain largely the sub-microscopic paste found in gouge. With the increased normal stress, high pressure metamorphism may occur and mylonitic textures form, destroying the clay minerals. In an area of hydrothermal flow, the fault breccia may be welded with calcite or quartz and the fault effectively healed until a stress buildup causes the welded material to fail. Fluids may exist over the width of the brecciated zone and active flow between part of the fault zone and surrounding rocks may be occurring at almost any scale. Fluid pressures in the whole width of the brecciated zone play a role in stress release on the active trace. Furthermore, any fault that can be demonstrated to be electrically conductive into the mid-crust or deeper probably

has significant fluid pressures associated with it. It remains to be seen whether through-crust, conductive faults can be mapped in detail.

Some faults are hypothesized to be associated with suture zones or overthrusts that involve fluid-rich sedimentary rocks. The electrical methods outlined can be used to study possibilities for the various geometries of faults that lie in ill-defined tectonic settings, using the fact that fluid-rich rocks provide easily mappable low-resistivity zones.

## SYNTHETIC MODELS

To understand the resolution of seismic and electrical geophysical methods to fault geometries we have used model simulations. Narrow (< 1 km wide) LVZ associated with faults have not been imaged unless they have very strong velocity reductions. Ellsworth et al. (1991) did a synthetic case with a broad LVZ having a sharp gradient at the fault (Figure 1a), using a typical distribution of earthquakes and seismic stations. The broad LVZ was imaged by the 3-D velocity inversion and the location of the fault was imaged as the place of maximum horizontal gradient, but the inversion was too coarse to image a sharp gradient. Note that the inversion model does a progressively poorer job with depth as it reverts to the initial vertically 1-D model below 16-km depth, while the true model continues with depth to be a laterally 1-D model. This weakness in the inversion model effects the hypocenters. To evaluate this a different set of hypocenters was used. The synthetic events have true locations on planes parallel to the fault. These events were located using the 3-D inversion model in Figure 1b in a joint inversion for hypocenters and station corrections. The relocated hypocenters tend to lie southwest of the true fault discontinuity (Figure 1c). The fairly accurate shallow hypocenters are 0.2 km southwest. There is a progressively greater mislocation with depth such that the relocated seismicity erroneously implies a steeply southwest-dipping fault rather than the actual vertical fault prescribed by the artificial data.

In an active source 2-D inversion study of the Columbia Plateau, Lutter et al. (1993) imaged LVZ fault zones within the basalts. The LVZ were 4.5 to 6 km wide at ≈ 2 km depth. He computed the image blurring inherent in the inversion, shown as ellipses in Figure 2. Removing the image blurring gives a estimate of the true width of the LVZ fault zones of 2 to 2.5 km. The superimposed ellipses indicate the horizontal and vertical components of image blurring for model parameters at the center of each ellipse. Comparing models A and B shows the effect of damping and gives some idea of the inherent nonuniqueness of inversion models. A has less damping than B. So A shows narrower deeper fault zones but has other small features that may be artifacts.

While the synthetic example above tried to represent a typical earthquake data set, for this Redbook workshop we constructed an example that would show the best possible results that we could hope to obtain. The model is a synthetic 2-D example with a narrow LVZ, but with dense station coverage and ideal hypocenter distribution. The true $V_p$ and $V_p/V_s$ models are shown in Figure 3. There are 90 earthquakes in the fault zone from 0.3 to 10 km depth, and 9 surface sources in the fault zone and at ± 15, 30, 45 km from the fault; and there are 74 stations. This exercise shows how well local earthquake tomography can image a narrow LVZ in an ideal case.

Inversion results done for $V_p$ only are shown in Figure 4. In these cases the earthquake hypocenters and origin times were treated as unknowns that were also solved for in the inversion, while the surface sources were treated as shots with known hypocenters and origin times. The coarsely gridded inversion model (Figure 4a) images the location of the LVZ but shows it as a broad feature with only a moderate velocity reduction. (Note that inversions were first done for a 1D initial model.) Finely gridded inversions included fine grid spacing over a 4-km wide region that includes the LVZ. When the fine-grid inversion is done with a 1D initial model (Figure 4b), it images a complicated LVZ. The LVZ looks simple in the upper few km, but with depth it broadens, splays and dips toward the high velocity block.

When the fine-grid inversion used the coarse inversion model (Figure 4a) as the initial model the LVZ was imaged much better (Figure 4c). The width of the LVZ is equivalent to the true width and it does not broaden with depth. Even though the amplitude of the imaged velocity reduction decreases with depth, the width of the LVZ at 9 km depth is the true width.

A simultaneous inversion for $V_p$ and $V_p/V_s$ was done using the coarse model as the initial model. All the events were treated as shots. The results are shown in Figure 5. The $V_p$ results for the $V_p$ only-with-shots inversion were equivalent to $V_p$ from the $V_p$ and $V_p/V_s$ inversion, and so are not presented. Both the width and the amplitude of the velocity reduction are imaged very well at all depths. It is quite evident that having known source locations at depth greatly improves the fault zone image. Uncertainty in the hypocenters weakens the imaged velocity contrast. The $V_p/V_s$ image is just as good as the $V_p$ image. This means that if we could obtain very high quality S-arrival times, we should be able to determine both $V_p$ and $V_p/V_s$ and so have more information with which to infer material properties and rheological behavior. Some artifacts are apparent, particularly at shallow depths. Such artifacts could be avoided or greatly reduced with higher damping as in the Columbia basalt example discussed above.

To provide a feeling for the resolution of inversions with electrical geophysical methods to image typical fault geometries, we used a model similar to that for the seismic simulation of Figure 1 and developed a synthetic MT inversion. The model consists of a 1000 ohm-m half-space (simulating granitic or metamorphic rocks), with a 5 km wide fault zone of 10 ohm-m, indicated by the dashed box B in Figure 6, (simulating breccia, gouge, fractures, and fluids). A shallow, conductive feature (3 ohm-m) on the west end of the profile simulates geological noise. Synthetic MT sounding locations are indicated by the triangles. The computed data for this model was then fed to a 2D inversion algorithm (Smith and Booker, 1991) to derive an image of the subsurface based upon inversion of the synthetic data. The inversion results are contoured in Figure 6. Like the seismic inversion, we obtain a very smoothed image of the actual fault zone. The width of the fault zone and its depth extent is reasonably well resolved by the resistivities lower than 300 ohm-m. However, it is also apparent that most of the response of the fault zone occurs in the upper 5 km, so that portions of a fault deeper than this would be difficult to associate with anything close to the true resistivities. The asymmetry of the response is due partially to the asymmetric station spacing and partially to the inversion trying to simultaneously resolve both the shallow feature, A, and the fault zone, B. This particular inversion tries to derive a minimally complicated structure using intermediate steps that are 1D approximations and we believe that other inversion methods that typically are unconstrained regarding model complexity might provide a closer image to the starting model. The fits of the "computed" to the "observed" data with this algorithm were not totally satisfactory, and as is normally the case, a series of forward modeling steps would normally be used to fine-tune the fit between the computed response and the observed response. Similar inversions could be made of DC, CSAMT, or other EM data for shallower, but higher resolution images of the fault zone. Such higher resolution images might be better for distinguishing fluids in the fault, but would be limited to depths above the seismogenic zone.

The model in figure 6 does not account for regional conductivity structures such as a nearby the ocean-continent margin. Including a realistic ocean-continent margin to the previous MT models causes the low frequency MT data to be highly affected, resulting in some loss of sensitivity to the fault zone. However, the effect of highly concentrated electrical currents in a nearby ocean can be used for study of the depth limits of the fault zone if we add a conductive mantle and assume that the fault extends through the crust into electrical contact with the mantle. This effect has been exploited by Park and others (1991) and Mackie and others (1993) to develop a hypothesis that the San Andreas fault extends through the crust. However, the differences in their models that resolve the through-crust fault are very small and occur only in the phase at the lower frequencies, thus the resolution concerning this connection and the resistivity of the deeper parts of the fault zone (they assume 300 ohm-m) is poor. In spite of the

difficulties in resolving such a crustal-cutting fault, the topic is very important to understanding deep fluids in faults such as the San Andreas and is an important area of research.

Electrical methods can also be used to image fluids systems that are not necessarily within the fault but in contact with it so that increases in fluid pressures can affect the fault. At Parkfield, California, on the San Andreas, the joint interpretation of local earthquake tomograms and MT inversions indicate a region of low-velocity and low-resistivity under Middle Mountain where the preparation zone for the characteristic cyclical earthquakes occurs. Eberhart-Phillips and Michael (1993) have interpreted that this zone at depths of 5-8 km represents high-pore-pressure brines. We suggest that the fluids may be related to oil field brines that have been forced under pressure from the Coalinga anticline region to Middle Mountain. Such a fluid model might help explain the characteristic nature of the earthquakes under Middle Mountain and new MT surveys are currently being interpreted to investigate this assumption.

## INTERPRETATION

Even if we have geophysically imaged a narrow zone with anomalous material properties, we still have a problem of interpreting the effect of fluids. Velocity and Poisson's ratio will be influenced by rock type, porosity and fractures, clay content, confining pressure and pore pressure, saturation and type of fluid. In a fault zone, gouge, breccia, fractures and pore pressure are all factors which are likely to be present and would contribute to low velocity. High conductivity is associated with fluids and very high conductivity is associated with saline fluids. However in a gouge, the clays have large surface areas that contain large amounts of bound water. Thus fluids imaged by electrical methods may not necessarily be free to move about, but may simply be an indicator of fault gouge, although some sensitivity to fluid salinity can be expected.

The figures below show one example of how difficult it is to separate out changes in effective pressure from changes in porosity ( $\Phi$ ) and lithology. An empirical relationship among seismic velocity, effective pressure ($P_e$ = confining pressure reduced by pore pressure), porosity, and clay content (C) in sandstone has been determined by Eberhart-Phillips, et al. (1989). Figure 7a shows the measured $V_p$ versus $V_p$ predicted by their relation.

$$V_p = 5.77 - 6.94\,\Phi - 1.73\ \text{sqrt}\ C + 0.446\,[\,P_e - 1.0\,e^{(-16.7\,P_e)}\,] \qquad (1)$$

For any particular velocity, there is a large set of $\Phi$, C, and $P_e$ that satisfy the relation, as illustrated by the heavy curves in Figure 8. There are several trends of outliers in Figure 7a, which represent rock samples that are not well fit by the empirical relation between $V_p$ and effective pressure ($P_e$ ). Thus there is an broader set of reasonable values of $P_e$, $\Phi$, clay content. To evaluate this additional uncertainty, we can consider the standard errors of the coefficients in (1), which was computed by stepwise multiple regression. These wider bounds are shown in Figure 8. Having both $V_p$ and $V_p/V_s$ would help reduce the uncertainty in estimating $P_e$.

Even though any $P_e$ relationship is too inexact to estimate small pore pressure changes, we may be able to use velocity to discern a change from hydrostatic to lithostatic pore pressure. Figure 7b shows the relationship between $V_p$ and $P_e$ in (1) with $\Phi$ and C fixed. Consider the case of a given rock unit at 5 km depth. $P_e$ is 0.66 kbar for hydrostatic pore pressure (point C in Figure 7b). If a $V_p$ decrease of 0.4 km/s is observed, the estimated $P_e$ would be 0.08 kbar, or near-lithostatic pore pressure (point D in Figure 7b). Such a large $\Delta V_p$ would probably be observable. The difficulty would be in deciding it was not caused by lithology. Having both $V_p$ and $V_p/V_s$ would help reduce the ambiguity in estimating $P_e$.

## SUMMARY AND CONCLUSIONS

The role of fluids in fault dynamics is extremely important, but the location, quantity, and pressures of these fluids is difficult to ascertain. We have described several geophysical methods which have been, or may be used, to address these problems. Faults are not always simple features, but represent broad zones with varying degrees of brecciation, fracturing, recrystalization, and alteration. It has been proven to be quite straightforward to map the low-velocity, low-resistivity breadth of active faults zones, but healed faults may not appear so anomalous. The more difficult problems is to sort out the actual fluid features of the fault zone from the other anomalous characteristics of the zone. Seismic tomography using local earthquakes can provide detailed P-wave and S-wave images that can be studies to indicate possible zone of fluids under pressure. Electrical sounding methods can similarly be used to locate fluid-rich regions in the fault, if they contain high proportions of dissolved solids so that they provide a more conductive environment than the remainder of the gouge zone. A more tractable geophysical problem is presented by the need to study fluid reservoirs in rocks lateral to the fault of interest; in this instance, it is quite possible to analyze the zone of permeable rocks that may play a major role in transmitting high-fluid-pressures to the fault. However, self-potential surveys have been used effectively to monitor transient changes in electrical field due to fluid motion within fault zones.

Future research into the problem of locating fluid-rich zones in faults from the surface probably should be focused on an integrated approach to the problem. Use of high-resolution seismic tomography with accurate ray-tracing algorithms, and waveform analysis should be used where local earthquakes are dense enough to sample the zones of interest. The seismic velocity and attenuation images can be combined with electrical geophysical images of the fault to provide clues to the fluid state of the target zones. In addition, electrical, gravity, and magnetic geophysical data can provide the needed framework to construct images of known shallow structures and interpret rock type in the volume of material surrounding the fault zone.

## REFERENCES


Ben-Zion, Y., S. Katz, and P. Leary, 1992, Joint Inversion of Fault Zone Head Waves and Direct P Arrivals for Crustal Structure Near Major Faults, J. Geophys. Res., 97, 1943.

Cormier, V. F., and P. Spudich, 1984, Amplification of ground motion and waveform complexity in fault zones: examples from the San Andreas and Calaveras faults, Geophys. J. Roy. Astr. Soc., 79, 135-152.

Corwin, R. F. and H. F. Morrison, 1977, Self-potential variations preceding earthquakes in central California, Geophysical Research Letters, 4, 171-174.

Diehl, S. F., and F. A. McKeown, 1991 Evidence for paleo- and recent fluid overpressure in the New Madrid seismic zone of the Reelfoot rift, southeast and northeast Arkansas (abs), Seism. Res. Letters, 62, 168-169.

Eberhart-Phillips, D., D.-H. Han, and M. D. Zoback, 1989., Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone, Geophysics, 54, 82-89

Eberhart-Phillips, D., and A. J. Michael, 1993, Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California, J. Geophys. Res., in press.

Ellsworth, W. L., D. M. Eberhart-Phillips, and E. Kissling, 1991, A test of local earthquake tomography, Seismol. Res. Lett., 62, 31.

Fitterman, D. V., 1979, Calculations of self-potential anomalies near vertical contacts, Geophysics, 44, 195-205.

Fung-Ming, C., 1976, An outline of prediction and forecast of Haicheng earthquake of M=7.3, Proceedings of the Lectures of the Seismological Delegation of the People's Republic of China, edited by P. M. Muller, Jet Propulsion Laboratory, Pasadena, CA.

Guo, H., A. Lerner-Lam, W. Caress, and S. Hough, 1992, Anomalous body wave attenuation in the mainshock region of the 1989 Loma Prieta, California earthquake, EOS, Trans. Am. Geophys. U., 73, 368.

Hamilton, R. M., and F. A. McKeown, 1988, Structure of the Blytheville arch in the New Madrid seismic zone, Seism. Res. Letters, 59, 117-121.

Hough, S. E., L. Seeber, A. Lerner-Lam, J. C. Armbruster, and H. Guo, 1991, Empirical Green's function analysis of Loma Prieta aftershocks, Bull. Seismol. Soc. Am., 81, 1737-1753.

Lin, C. H., and S. W. Roecker, 1993, Effects of ray tracing on the determination of earthquake hypocenters, focal mechanisms, and velocity structures in the Morgan Hill / Coyote Lake area of central California, unpublished research report.

Lin, C. H., and S. W. Roecker, 1993, Hypocenters, focal mechanisms, and three-dimensional velocity structure of the Bear Valley area of central California, unpublished research report.

Lutter, W. J., R. D. Catchings, and C. M. Jarchow, 1993, An image of the Colombia Plateau from inversion of high-resolution seismic data, Geophysics, submitted.

Mackie, Randall L., T. R. Madden, and Edward A. Nichols, 1993, A magnetotelluric survey of the Loma Prieta earthquake area: implications for earthquake processes and lower crustal conductivity. (in press) U.S. Geological Survey Professional Paper.

Michelini, A., and T. V. McEvilly, 1991, Seismological studies at Parkfield: I. Simultaneous inversion for velocity structure and hypocenters using cubic b-splines parameterization, Bull. Seismol. Soc. Am., 81, 524-552.

Miyashiro, A., 1973, Metamorphism and metamorphic belts. George Allen & Unwin, London, 492 p.

Mooney, W. D., and T. M. Brocher, 1987, Coincident seismic reflection / refractions studies of the continental lithosphere: a global review, Rev. Geophys., 25, 723-742.

Murakami, Hideki, Hitoshi Mizutani, and Sachio Nabetani, 1984, Self-potential anomalies associated with an active fault, Journal of Geomagnetism and Geoelectricity, 36, 351-376.

Park, S. K., G. P. Biasi, R. L. Mackie, and T. R. Madden, 1991, Magnetotelluric evidence for crustal suture zones bounding the southern Great Valley, California: Journal of Geophysical Research, 96, 353-376.

Rodriguez, B. D., V. F. Labson, W. D. Stanley, and J. A. Sampson, 1993, Deep geoelectric structure of the Loma Prieta, CA earthquake region: U.S. Geological Survey Professional Paper, submitted.

Scott, J. S., 1992, Microearthquake studies in the Anza seismic gap, Ph.D. Thesis, 277 pp., Univ. Calif. San Diego.

Sobolev, G. A., V. N. Morozov, and N. I. Migunov, 1972, The electrotelluric field and a strong earthquake on Kamchatka, Isvestia, Earth Physics, 2, 73-80.

Smith, J. T., and J. R. Booker, 1991, Rapid inversion of two and three-dimensional magnetotelluric data, J. Geophys. Res., 96, 3905-3922.

Zucca, J. J., L. J. Hutchings, and P. W. Kasameyer, 1993, Velocity and attenuation structure of The Geysers geothermal field, California, Law. Liv. Nat. Lab. UCRL-JC-113393.

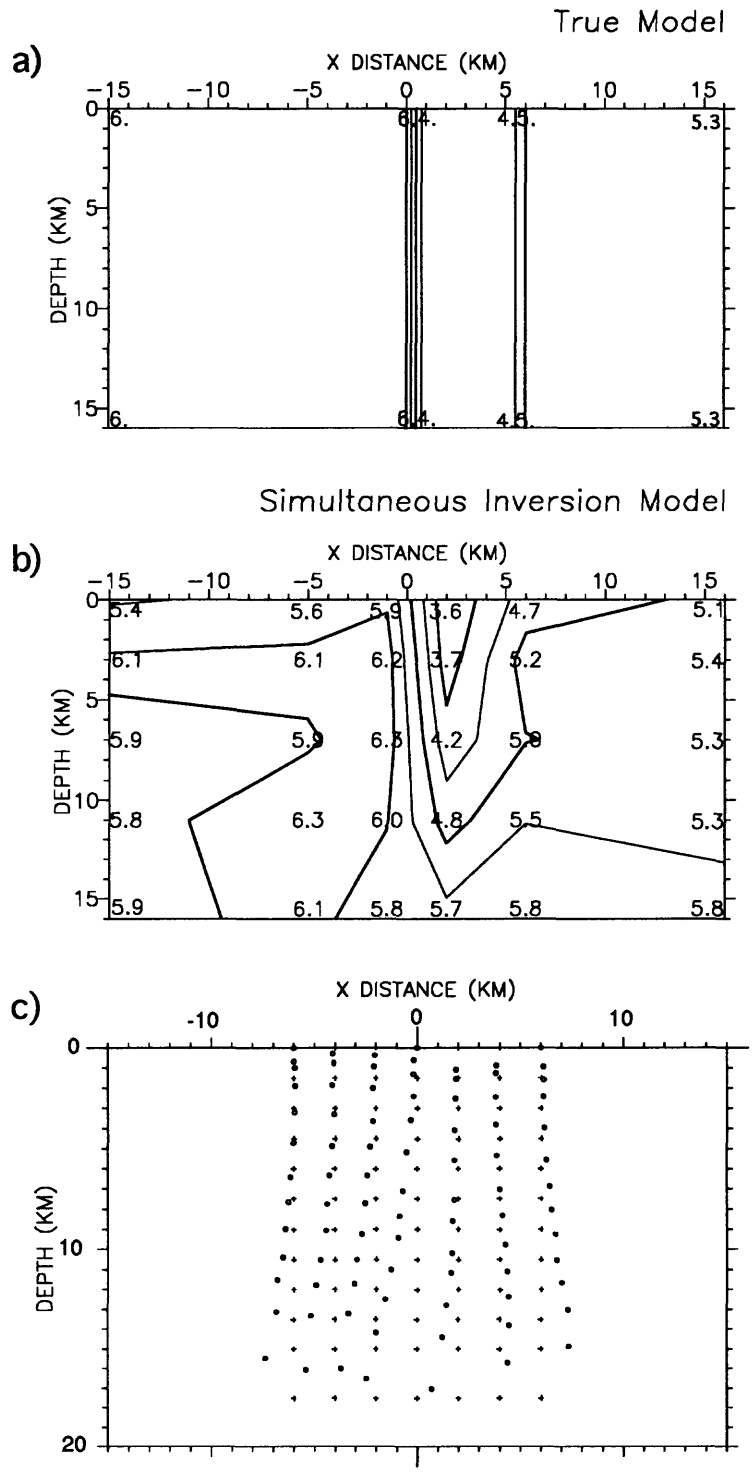



Figure 1. Example of the simultaneous inversion method used on a synthetic data set (typical distribution of earthquakes and seismic stations) that was generated for a laterally-varying 1-D test model. Velocity is labelled at gridpoints and contoured at 0.5-km/sec intervals. a) Cross-section of the true model. b) Cross-section of simultaneous 3-D inversion model. c) Test of relocating synthetic hypocenters: pluses show true locations, circles show relocations obtained using model (b) in a joint inversion for hypocenters and station corrections. (Note that a different set of hypocenters was used to obtain (b), which was much more unevenly distributed). [from Ellsworth, et al., 1991.]

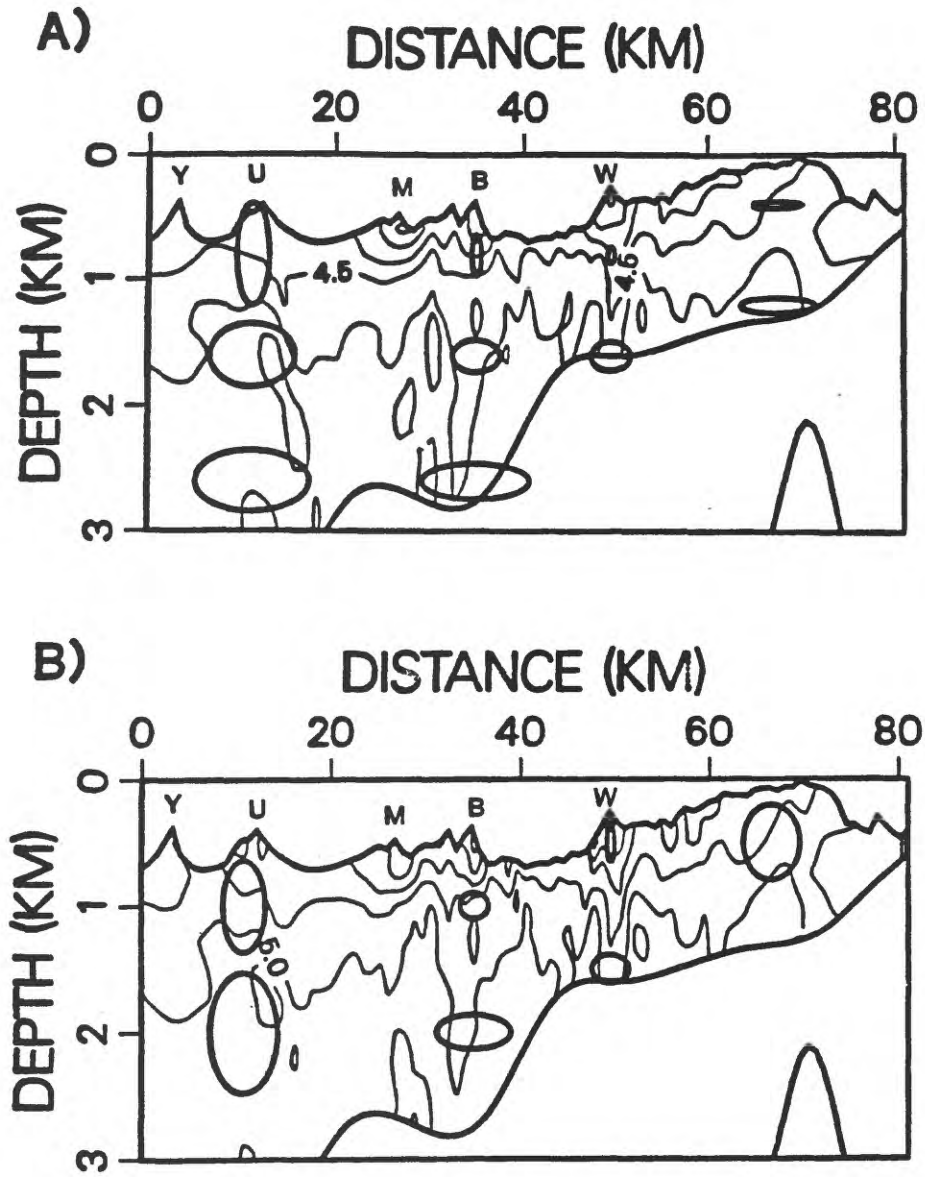



Figure 2. Upper 3 km of 2-D inversion model with basalt velocity structure contoured. Superimposed ellipses indicate the horizontal and vertical components of image blurring for model parameters at the center of each ellipse. (B) has higher damping than (A) [from Lutter, W. J., R. D. Catchings, C. M. Jarchow, An image of the Columbia Plateau from inversion of high-resolution seismic data, Geophysics, sub. 1992]

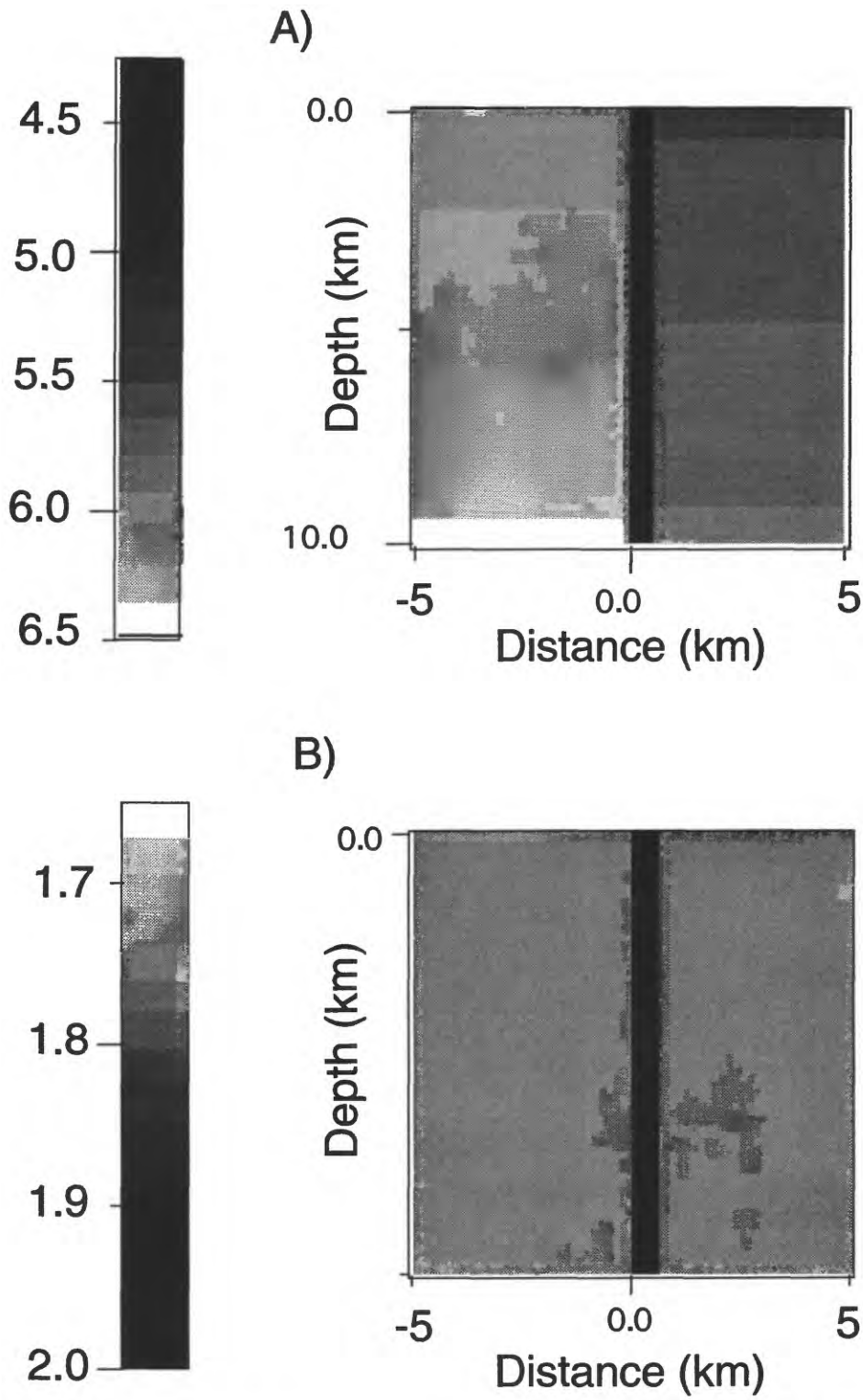



Figure 3. Synthetic fault zone example with narrow (< 1 km wide) (A) low Vp and (B) high Vp/Vs was studied to evaluate an ideal case. Hypocenters were along each side of the zone from 0 to 10 km depth, also surface sources at ±15, ±30, ±45 km horizontal distance.

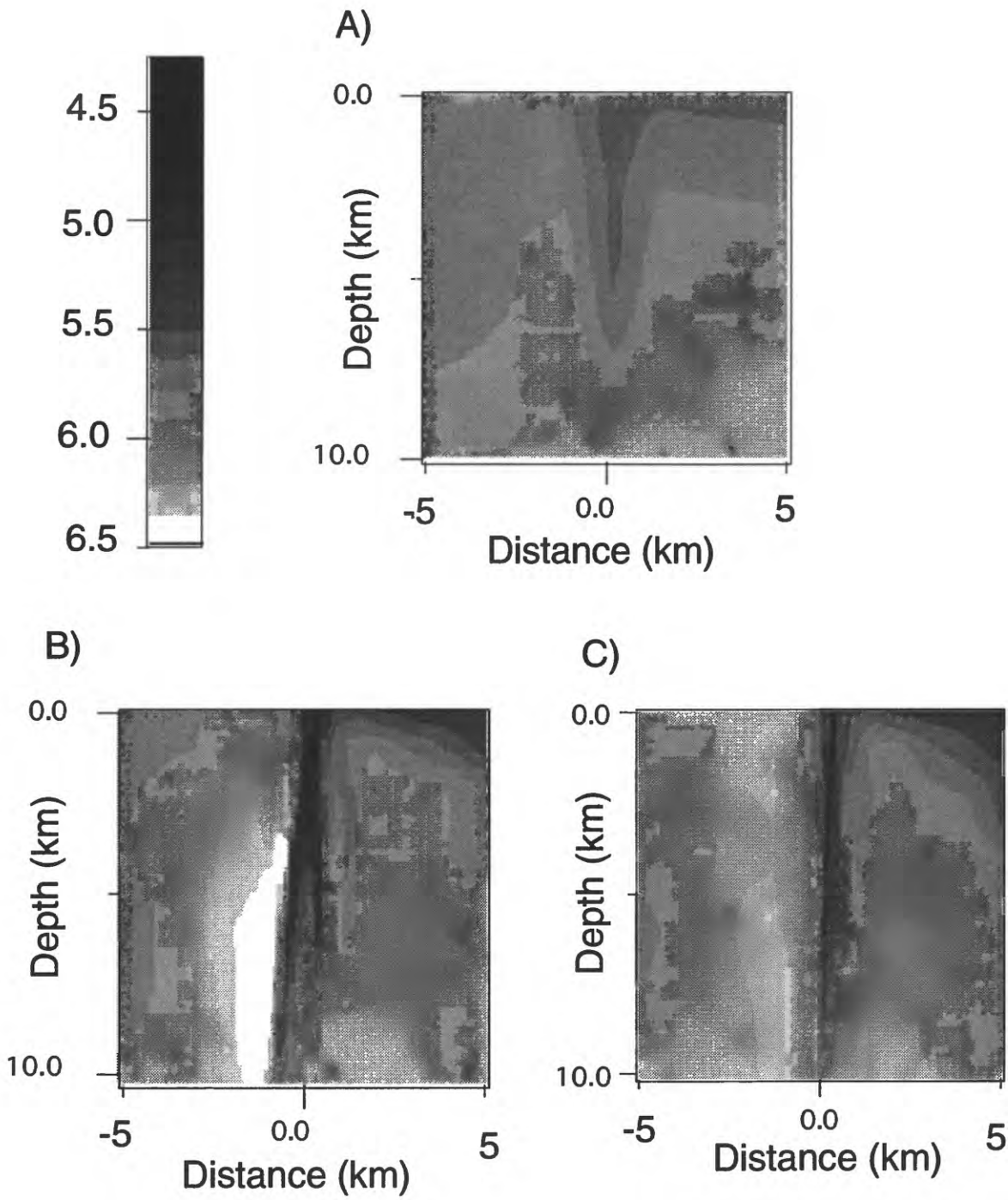



Figure 4. Vp inversion results from simultaneous inversions for velocity and hypocenters. A) Coarse grid, 1D initial model. B) Fine grid, 1D initial model. C) Fine grid, coarse grid result used as initial model.

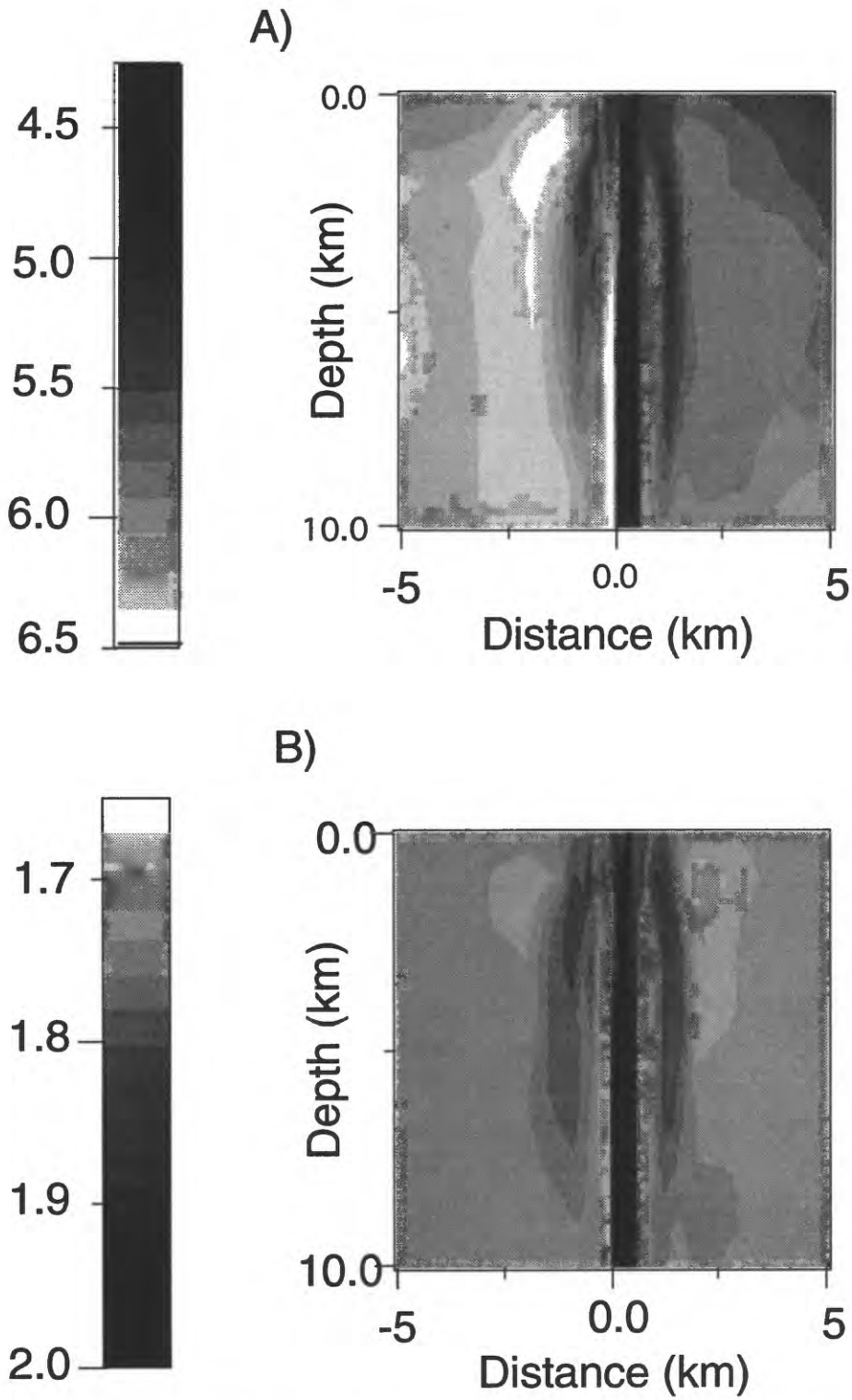



Figure 5. Inversion results using all synthetic events as shots (known hypocenters and origin times) for (A) Vp and (B) Vp/Vs.

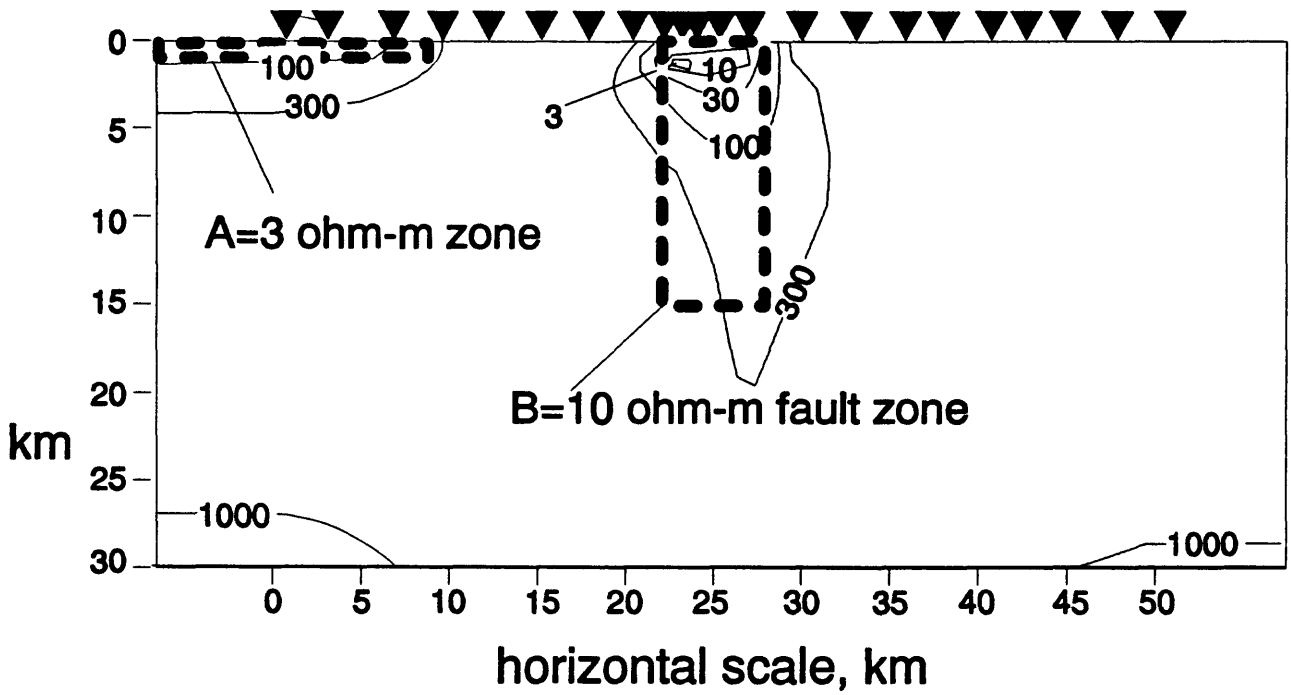
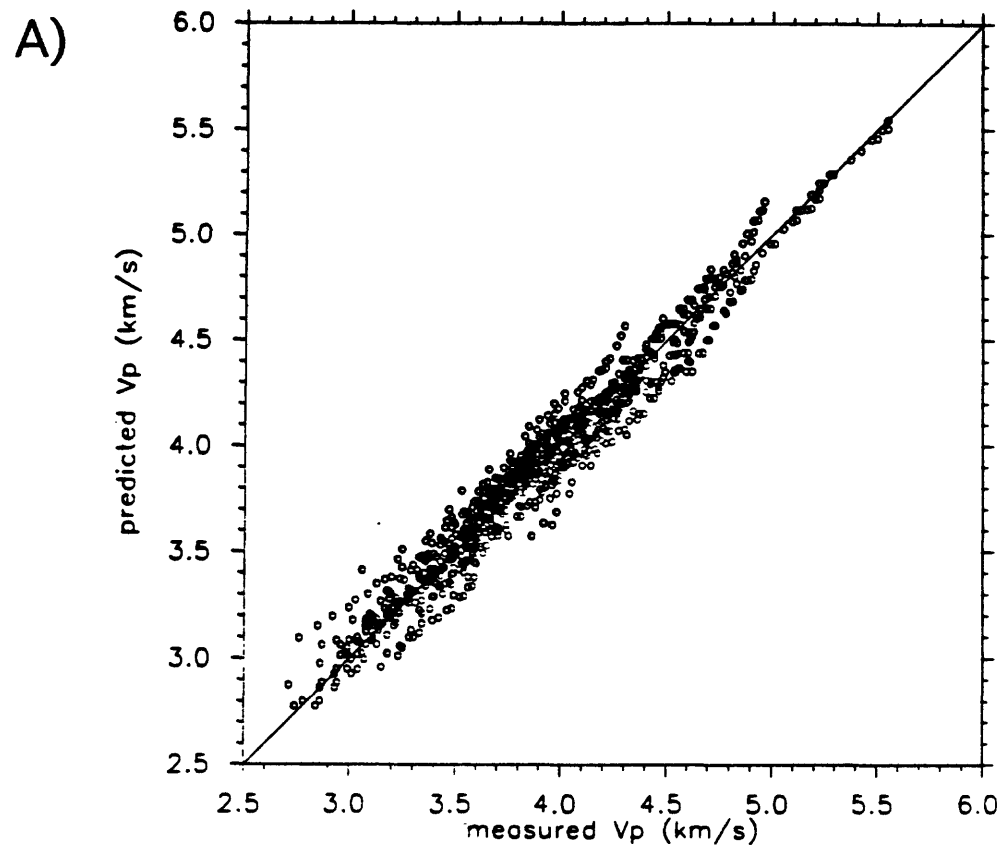



Figure 6. Inversion of synthetic MT data for fault model similar to that used in the seismic simulation (Fig. 3). Vertical dotted box (B) is the 5 km wide fault zone with a resistivity 10 ohm-m and horizontal dotted box (A) is a surface feature of 10 ohm that serves as geologic "noise". Background resistivity outside of the two lower resistivity boxes is 1000 ohm-m. Inversion results inresistivity are shown as contours at logarithmic intervals, in ohm-m. Inverted triangles are MT measurement points.



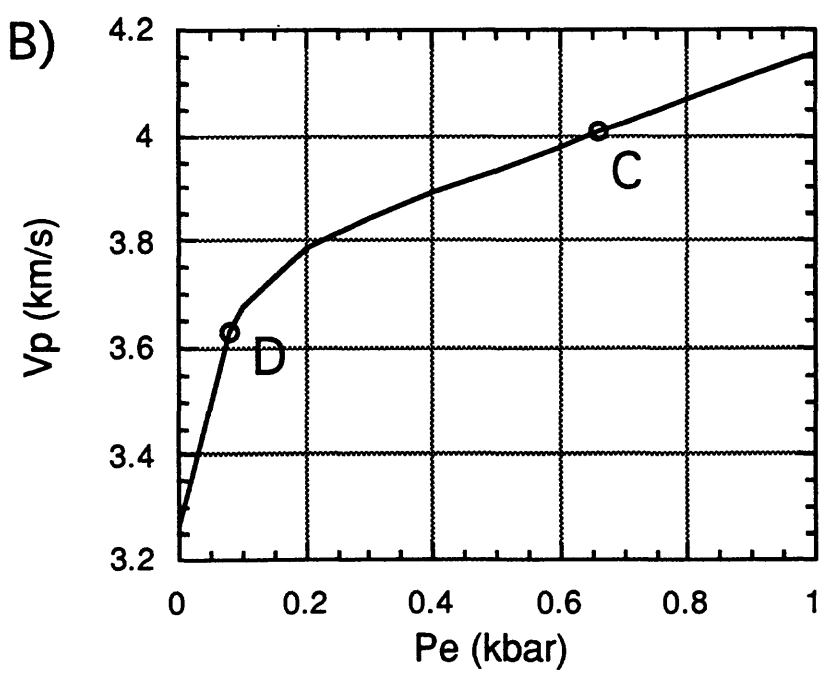




Figure 7. A) Measured Vp versus predicted Vp, calculated using equation (1). Line represents perfect fit.
B) Relationship between Vp and Pe (confining pressure reduced by fluid pressure) from equation (1). Note that the minimum value of Vp (at Pe = 0) depends on porosity and clay content. Point C represents case of hydrostatic fluid pressure at 5 km depth. Point D represents an increase in fluid pressure to near-lithostatic corresponding to a velocity reduction of 0.4 km/s.

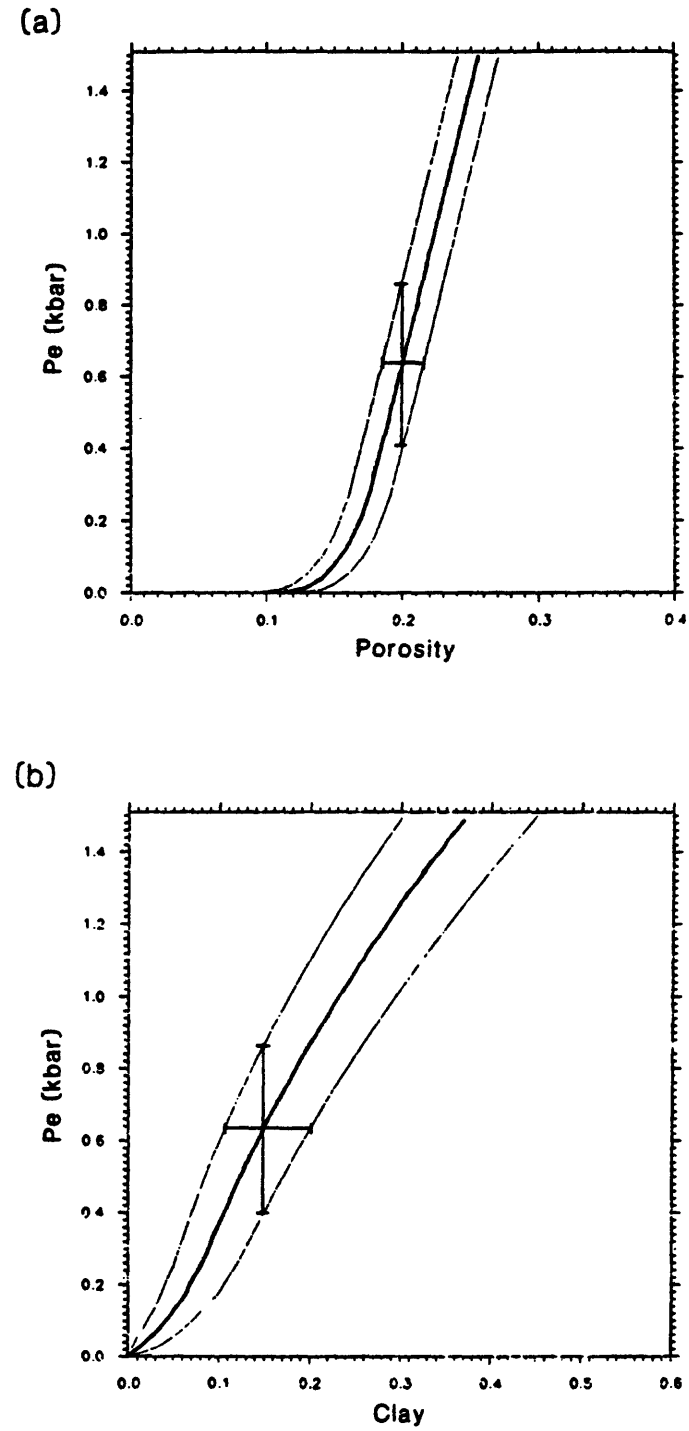



Figure 8. a) Curves showing $V_p$ = 4.00 bounded by curves for 1 SE. a) Shown in $\phi$–$P_e$ plane for C=0.15. b) Shown in C-$P_e$ plane for $\phi$=0.20. Uncertainties in individual parameters for 1 SE are indicated for the point ($\phi$=0.20,C=0.15,$P_e$=0.62).

# FLUIDS IN THE EARTH'S CRUST:
## Electromagnetic Inferences on Existence and Distribution

Philip E. Wannamaker
University of Utah Research Institute
391 Chipeta Way, Suite C
Salt Lake City, UT 84108-1295

Over the last decade, we have witnessed striking advances in the ability of electromagnetic (EM) geophysics to image resistivity structure of the crust. One of the best known capabilities of electrical resistivity research in Earth science is the potential to infer the presence of fluids in the crust. Less often is appreciated the complexity of the inferred fluids and the general need to interpret resistivity structure in concert with external geoscientific constraints. However, despite problems with the accuracy of early resistivity models, a positive correlation between temperature and low resistivity appears to be justified for the deep crust. This may be considered as evidence for a fluid cause of low resistivity, although the composition of any such fluid must be a complex function of the physical state. Research still is needed on the stability and properties of fluids at high metamorphic grades and on the importance of non-fluid mechanisms of reducing resistivity. Nevertheless, modern EM surveys integrated with other methods have augmented models of fluid existence and distribution. For example, injection of fluids to great depth in subduction environments may result in massive retrograde metamorphism of wide areas of the lower lithosphere. Low-resistivity anomalies observed even in Precambrian suture environments suggest that low resistivity of fluids or former sedimentary material perhaps can be preserved indefinitely barring thorough prograde metamorphic overprinting. The role of retrograded collision zones in controlling subsequent deformation remains an outstanding question. Under special structural circumstances, EM responses can be very sensitive to crustal electrical windows which possibly correspond to pathways through fault zones connecting upper and lower crustal fluids. This phenomenon cannot be observed everywhere, but careful experiments in appropriate 'type' environments carry the promise of unprecedented resolution using EM methods.

## Introduction

One of the most widely recognized roles for electrical resistivity research in Earth science is the capability to infer the domains and pathways of fluids (including melts) in the crust and mantle, and their attendant structures (e.g., Wannamaker and Hohmann, 1991; Jones, 1992). By fluids, I mean not only free fluids developed in or moving through the Earth in situ, but also fluid-bearing earth materials transported tectonically such as in subduction zones or overthrust environments. Achieving this capability has required developments both in the relation between resistivity and Earth conditions, and in specific technology to image electrical properties.

Prior to the 1980's, electrical resistivity was thought to be able to imply deep temperatures in the crust by comparing Earth resistivity profiles derived from field electromagnetic (EM) measurements to laboratory measurements of the solid-state resistivity of dry, representative crustal minerals. As field models accumulated, it became apparent that

temperatures so implied appeared systematically higher (by 50 to 100%) than those predicted by geotherms derived from essentially coincident heat flow data (Shankland and Ander, 1983; Figure 1). It is concluded usually that solid-state (semi-conduction) mechanisms of resistivity are not completely representative of those of the deep crust, and that an additional, low-resistivity boundary phase is necessary. Due to its common occurence and property as a universal solvent, this boundary phase typically is interpreted to be free water (e.g., Hyndman and Shearer, 1989).

## Conditions of Fluid Existence and Competing Mechanisms

One of the long standing problems with free water as a cause of reduced resistivity in the lower crust has been the inference of high metamorphic grade there from studies of xenoliths and exhumed terranes, in which free water at lithostatic pressure should react with ambient silicate mineralogy to form hydrate minerals which do not enhance electrical conduction (Yardley, 1986; Wannamaker, 1986). Also, recent seismic modeling shows that impedance contrasts suitable for generating observed, sub-horizontal reflectors in the deep crust may be created by mafic/ultramafic layered intrusions (Warner, 1990; Singh and McKenzie, 1993). However, a mixed $H_2O$-$CO_2$ fluid is more typical in composition here and can exist in equilibrium at much lower temperatures due to reduced activity of $H_2O$ (Figure 2). Despite some claims, such a mixed fluid is probably of high ionic strength (Wannamaker, 1986; Nesbitt, 1993) and a good electrical conductor. Very high salinity also can lower further the temperature of fluid-rock reaction. Aqueous fluids are stable in rocks of lower metamorphic grade, and in most rocks in the brittle mechanical regime (Wannamaker, 1986; Fournier, 1991), the latter of which correspond to temperatures less than about 400°C and in shield areas may exist to depths beyond 30 km (e.g., Kellett et al., 1992). Interconnectivity of $H_2O$-$CO_2$ fluids under ductile, lower crustal conditions is of equal importance to thermodynamic stability, and experimental evidence so far does not indicate that interconnectivity readily occurs (Watson and Brenan, 1987; Watson and Lupulescu, 1993). However, fluid interconnection seems possible at least in a small-scale fracture network at modest fluid overpressures (Ferry, 1987; Brenan and Watson, 1988) and residence times of such a fluid may be substantial under deep 'aquitards' represented by metamorphic or igneous foliations (Thompson and Connelly, 1990). Finally, Peterson and Newton (1989) show that $CO_2$-bearing lower crustal rocks with pyroxene or hydrous equivalents (a likely situation) will also melt at relatively low temperatures (around 700°C). The resulting melt appears very fluid and might readily interconnect.

More recently, regionally interconnected films of conductive graphite are proposed to explain low resistivity deep in the crust. It is argued that $CO_2$ fluids may enter the graphite stability regime if $fO_2$ is controlled by exsolving Fe-Ti oxides (Frost et al., 1989), which may define $fO_2$ along grain boundaries even if the bulk rock tends to lie closer to the quartz-fayalite-magnetite buffer. However, in addition to perceived uncertainties in connecting $CO_2$-bearing fluids over regional distances, graphite conductivity exhibits little temperature dependence (Duba et al., 1988), while the magnitude and shallowness of the lower crustal low resistivity zone evidently does (Shankland and Ander, 1983; Majorowicz et al., 1993; Figure 3). One must also distinguish between such metamorphic graphite and biogenic graphite in former sedimentary packages carried to depth in compressional tectonism, which is more widely accepted as an important cause of low resistivity structure in the crust (Wannamaker and Hohmann, 1991). Furthermore, solid-state mechanisms particular to feldspars (the crust is feldspar-dominated in

composition) have received inadequate attention. These defect-rich minerals may have resistivities which are orders of magnitude lower than pyroxenes or olivine but this is rarely recognized in the lab due to measurement times too short to allow sample equilibrium (Duba, 1976; Figure 4). Defect populations are enhanced in the presence of finite $P(H_2O)$, even if $P(H_2O)$ is too low to form a free fluid phase. There should be attempts to quantify this conductivity in measurements buffered by hydrate minerals representative of the lower crust (Wannamaker and Hohmann, 1991).

## Field Resistivity Studies and Fluid Implications

Illustrative field surveys here should provide a sense of the resolving capability of EM data both in terms of basic model geometry and of inferences about Earth conditions. Examples will appear from active and fossil regimes, and from compressional, extensional and stable environments.

The EMSLAB experiment was successful in delineating resistivity structure associated with subduction of the young Juan de Fuca plate beneath western North America (Wannamaker et al., 1989; Figure 5). This structure has been interpreted as providing a mechanism for the injection of water into the deep continental crust and for causing widespread retrograde metamorphism therein. Accretion and arc magmatism are viewed as important mechanisms of continental growth. The degree to which a hydrated crustal column is overprinted by prograde metamorphism during closure of subduction (cratonization) remains somewhat open to question, I believe. Numerous examples exist of pronounced low resistivity zones remaining even in suture zones of Precambrian age. Such zones extend to lower crustal depths, dip in the direction of former subduction, and are correlated with dipping packages of seismic reflectors similarly interpreted (e.g., Jones, 1992; Korja et al., 1993; Figure 6). The conductance of such zones can be so great as to suggest a mechanism in addition to fluids; graphite from pyrolysis of former sedimentary material is commonly invoked. Establishing the existence of such material and its role in controlling the loci of subsequent orogenies is a fundamental question in the evolution of the Earth.

Typical deep resistivity profiles from shield and active extensional environments are compared in Figure 7. The shield profile is from the early work of Dowling (1970) but more recent models by Kellett et al., (1992) and Kurtz et al., (1993) are very similar. The profile displays a modest resistivity minimum in the depth range 10 to 40 km roughly, and a further drop in resistivity below about 100 km (not shown). I interpret that crustal low-resistivity layer as due possibly to aqueous brines in fractures and pores in rocks of the brittle regime, which may extend to such depths. The extensional profile is from the eastern Great Basin province (Wannamaker and Johnston, 1991). This region of high heat flow and geothermal occurences exhibits a low-resistivity layer in a depth interval similar to that of the shield, within uncertainty, although its magnitude is greater. The upper portion of this layer is consistent with hydrothermal fluids below the brittle-ductile transition which are impeded by reduced permeability from ascending to higher levels (Stanley et al., 1990; Fournier, 1991; Jones, 1992). Depth extent independent of layer conductivity is difficult to resolve with EM, but the layer appears to encompass the crust in this region. Aqueous fluids in the lowermost crust are more difficult to justify, as discussed above, due to inferred granulite metamorphic grade. In this regime, $CO_2$-$H_2O$ fluids or induced melts, or perhaps defect-controlled solid-state mechanisms, must be called upon. While both the shield and extensional resistivity structures could be

described as due to 'brines', and most interpretations I have read appear to stop about there, the resistivity models clearly are capability of offering much more information about physicochemical state. An important point here is, that there is a plurality of causes of resistivity structure and this property should be interpreted within a framework of independent geoscientific constraints. Principally, these constraints include the geotherm and prior metamorphic history.

A recent and rather compelling case for the role of increased temperature and evolved fluids in the lower crust is provided by Booker et al. (1992) who acquired MT profiles over northern California both north and south of the subducted Gorda Plate edge (Figure 8). In concept, overlying crust of similar geologic history is subjected to different 'step' thermal inputs from below, with the southern profile experiencing an increase in temperature over a postulated half spreading center and the northern profile experiencing none. A pronounced middle to lower crustal conductor exists below the southern profile which is not so apparent below the northern one; it is proposed to represent fluids released when metamorphic rocks of the California Coast Ranges are heated by the upwelling asthenosphere south of the Gorda Plate. It is argued that these fluids may be responsible for a ductile lower crust which controls seismicity and tectonics of much of western California.

Finally, a recently-recognized phenomenon in EM induction is proving to be powerful in identifying potential fluid pathways between the upper and lower crust. Strong electrical current systems are induced in the conductive ocean water by natural EM (magnetotelluric or MT) fields, and continuity of these currents tends to be preserved as they are driven onto the land (Mackie et al., 1988). Taking the paths of least resistance, such currents may be abruptly channeled down more conductive fault zones to complete the current circuit with the less-resistive lower crust or upper mantle. Although it is difficult to obtain high quality EM data in areas of strong EM cultural interference, Mackie et al. (1993, in press) have observed such abrupt changes in the MT response for stations taken on either side of the San Andreas fault system in the vicinity of the Loma Prieta earthquake (Figure 9). Immediate followup studies are proposed over locked and creeping sections of the fault system to assess possible differences in fluid content. Such crustal current windows have been exploited as well to infer the position of Late Mesozoic or Early Cenozoic sutures in central California (Park et al., 1991).

Another application of the phenomenon is detection of possible deep transition boundary faults separating the Great Basin and Colorado Plateau provinces (Wannamaker and Johnston, 1991). A 2-D resistivity model for twenty MT stations covering the Milford Valley and Mineral Mountains to the foothills of the Tushar Mountains in S. W. Utah was derived by forward modeling and parameterized inversion using a two-dimensional finite element algorithm and emphasizing the transverse magnetic (TM) mode observations (Figure 10). The TM data from soundings straddling the possible projection of the Hurricane fault zone into study area show an abrupt change in impedance phase and slope of apparent resistivity, especially around 10 s period (Figure 11). This change implies a low-resistivity connection between the upper crust and lower crust in the area. The figure also indicates the poorer fit which results when such a pathway is removed and the best-fitting structure otherwise is sought. In this region, the 'ocean' of electric currents analogous to the Pacific Ocean in the San Andreas study may include the sedimentary section of the Colorado Plateau to the east. Future research involves extending the MT profile to the east to search for additional crustal low-resistivity windows. The Great Basin and San Andreas structures suggest that upper and lower crustal fluids can be connected; the degree of mixing possible over geological time is unknown.

## Conclusions

Over the last decade, we have witnessed striking advances in the ability of EM geophysics to image resistivity structure of the crust. Not quite so striking has been our understanding of the diverse collection of physico-chemical mechanisms which control resistivity, but this question is key in the value of this field to geoscience. The time should be past when deep resistivity profiles are used simply to estimate crustal temperatures. Merely assigning anomalous deep structure to aqueous brines gives short shrift to the potential of resistivity data to illuminate conditions in the Earth. Modern EM surveys integrated with other geoscientific methods have augmented models of fluid existence and distribution in the crust. Injection of fluids to great depth in subduction environments may result in massive retrograde metamorphism of wide areas of the lower lithosphere. Low-resistivity structures associated with such suture zones have persisted since the Precambrian and their role in controlling subsequent deformation remains an outstanding question. Despite problems with the accuracy of early resistivity models, a positive correlation between temperature and low resistivity appears to be justified for the deep crust. While this may be considered as evidence for a fluid cause of low resistivity, the composition of the fluid surely is a complex function of the physical state. Research still is needed on the stability and properties of fluids at high metamorphic grades and on the importance of non-fluid mechanisms of low resistivity. Under special structural circumstances, EM responses can be very sensitive to crustal electrical windows which possibly correspond to pathways through fault zones connecting upper and lower crustal fluids. This phenomenon cannot be observed everywhere, but careful experiments in type environments have the potential for unprecedented resolution using EM methods.

## References


Booker, J. R., C. Aprea, J. T. Smith, H. Waff, and G. Egbert, Magnetotelluric evidence for ductility in the lower crust of Cape Mendocino, EOS, Trans. AGU, 73(43), Supplement, 1992.

Brenan, J. M., and E. B. Watson, Fluids in the lithosphere, 2. Experimental constraints on $CO_2$ transport in dunite and quartzite at elevated P-T conditions with implications for mantle and crustal decarbonation processes, Earth Planet. Sci. Lett., 91, 141-158, 1988.

Dowling, F. L., Magnetotelluric measurements across the Wisconsin Arch, J. Geophys. Res., 75, 2683-2698, 1970.

Duba, A., Are laboratory electrical conductivity data relevant to the earth?, Acta Geod. Geophys. Montanist., Acad. Sci. Hung., 11, 485-495, 1976.

Duba, A., E. Huenges, G. Nover, G. Will, and H. Jodicke, Impedance of black shale from Munsterland 1 borehole: an anomalously good conductor?, Geophys. J., 94, 413-419, 1988.

Ferry, J. M., Metamorphic hydrology at 13-km depth and 400-500°C, Amer. Mineral., 72, 39-58, 1987.

Fournier, R. O., The transition from hydrostatic to greater than hydrostatic fluid pressure in presently active continental hydrothermal systems in crystalline rock, Geophys. Res. Lett., 18, 955-958, 1991.

Frost, B. R., W. S. Fyfe, K. Tazaki, and T. Chan, Grain-boundary graphite in rocks and implications for high electrical conductivity in the lower crust, Nature, 340, 134-136, 1989.

Hyndman, R. D., and P. M. Shearer, Water in the lower continental crust: modeling magnetotelluric and seismic reflection results, Geophys. J. Int., 98, 343-367, 1988.

Jones, A. G., Electrical conductivity of the continental lower crust, in Continental lower crust, ed. by D. M. Fountain, R. J. Arculus and R. W. Kay, Elsevier, 81-143, 1992.

Karato, S., The role of hydrogen in the electrical conductivity of the upper mantle, Nature, 347, 272-273, 1990.

Kellett, R. L., M. Mareschal, and R. D. Kurtz, A model of lower crustal electrical anisotropy for the Pontiac subprovince of the Canadian shield, Geophys. J. Int., 111, 141-150, 1992.

Korja, A., T. Korja, U. Luosto, P. Heikkinen, Seismic and Geoelectric Evidence for Collisional and Extensional Events in the Fennoscandian Shield - Implications for Precambrian Crustal Evolution, Tectonophysics, 219, 129-152, 1993.

Kurtz, R. D., J. A. Craven, E. R. Niblett, and R. A. Stevens, The conductivity of the crust and upper mantle beneath the Kapuskasing Uplift: electrical anisotropy in the upper mantle, Geophys. J. Int., 113, 483-498, 1993.

Mackie, R. L., B. R. Bennett, and T. R. Madden, Long-period magnetotelluric measurements near the central California coast: a land-locked view of the conductivity structure under the Pacific Ocean, Geophys. J. Int., 95, 181-194, 1988.

Mackie, R. L., T. R. Madden, and E. A. Nichols, A magnetotelluric survey of the Loma Prieta earthquake area: implications for earthquake processes and lower crustal conductivity, U. S. Geol. Surv., Prof. Paper on the Loma Prieta earthquake, 1993, in press.

Majorowicz, J. A., D. I. Gough, and T. J. Lewis, Electrical conductivity and temperature in the Canadian Cordilleran crust, Earth Planet. Sci. Lett., 115, 57-64, 1993.

Nesbitt, B. E., Electrical resistivities of crustal fluids, J. Geophys. Res., 98, 4301-4310, 1993.

Park, S. K., G. P. Biasi, R. L. Mackie, and T. R. Madden, Magnetotelluric evidence for crustal suture zones bounding the southern Great Valley, California, J. Geophys. Res., 96, 353-376, 1991.

Peterson, J. W., and R. C. Newton, $CO_2$-enhanced melting of biotite-bearing rocks at deep-crustal pressure-temperature conditions, Nature, 340, 378-380, 1989.

Shankland, T. J., and M. E. Ander, Electrical conductivity, temperature, and fluids in the lower crust, J. Geophys. Res., 88, 9475-9484, 1983.

Singh, S. C., and D. McKenzie, Layering in the lower crust, Geophys. J. Int., 113, 622-628, 1993.

Stanley, W. D., W. D. Mooney, and G. S. Fuis, Deep crustal structure of the Cascade Range and surrounding regions from seismic refraction and magnetotelluric data, J. Geophys. Res., 95, 19,419-19,468, 1990.

Thompson, A. B., and J. A. D. Connolly, Metamorphic fluids and anomalous porosities in the lower crust, Tectonophys., 182, 47-55, 1990.

Wannamaker, P. E., Electrical conductivity of water-undersaturated crustal melting, J. Geophys. Res., 91, 6321-6327, 1986.

Wannamaker, P. E., J. R. Booker, A. G. Jones, A. D. Chave, J. H. Filloux, H. S. Waff, and L. K. Law, Resistivity cross section through the Juan de Fuca subduction system and its tectonic implications, J. Geophys. Res., 94, 14,127-14,144, 1989.

Wannamaker, P. E., and G. W. Hohmann, Electromagnetic induction studies, U. S. National Rep. to IUGG, Rev. Geophys., in press, 1991.

Wannamaker, P. E., and J. M. Johnston, Uncovering Sevier overthrust structure and potential hydrocarbon source rock using magnetotellurics, Utah Geological Survey, Open File Report, 13 p., 1991.

Warner, M., Basalts, water, or shear zones in the lower continental crust?, Tectonophys., 173, 163-174, 1990.

Watson, E. B., and J. M. Brenan, Fluids in the lithosphere, 1. Experimentally determined wetting characteristics of $CO_2$-$H_2O$ fluids and their implications for fluid transport, host-rock physical properties, and fluid inclusion formation, Earth and Planet. Sci. Lett., 85, 497-515, 1987.

Watson, E. B., and A. Lupulescu, Aqueous fluid connectivity and chemical transport in clinopyroxene-rich rocks, Earth and Planet. Sci. Lett., 117, 279-294, 1993.

Yardley, B. W. D., Is there water in the deep continental crust?, Nature, 323, 111, 1986.

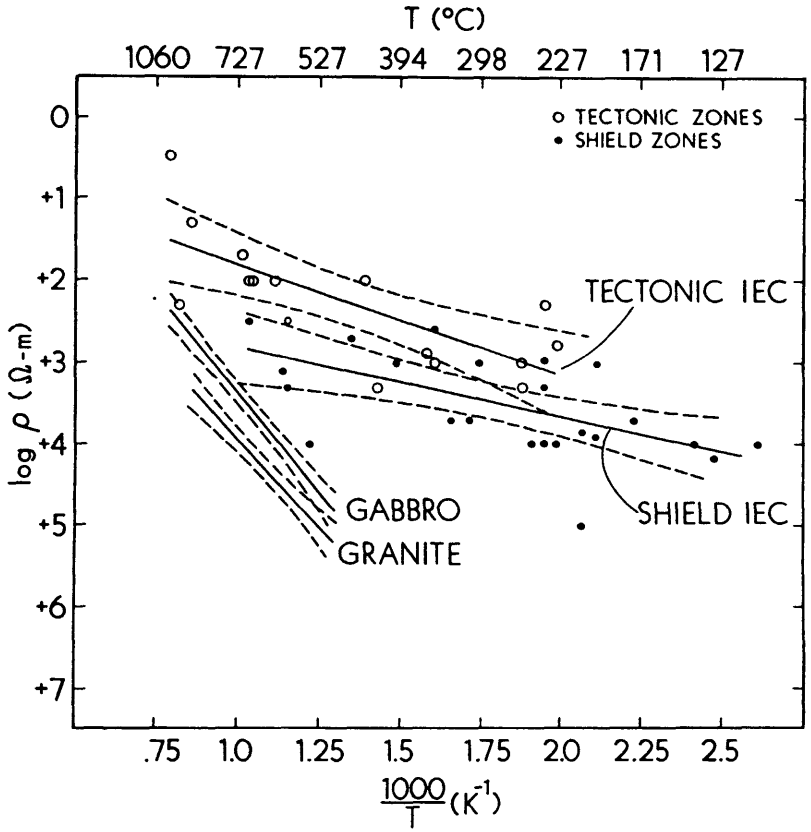



Figure 1. Logarithm of electrical conductivity versus reciprocal temperature for tabulated intra-crustal electrical conducting layers (IEC) in tectonic regions (circles) and stable regions (crosses). Included are dry solid-state conductivities for gabbros and granites. Modified from Shankland and Ander (1983).

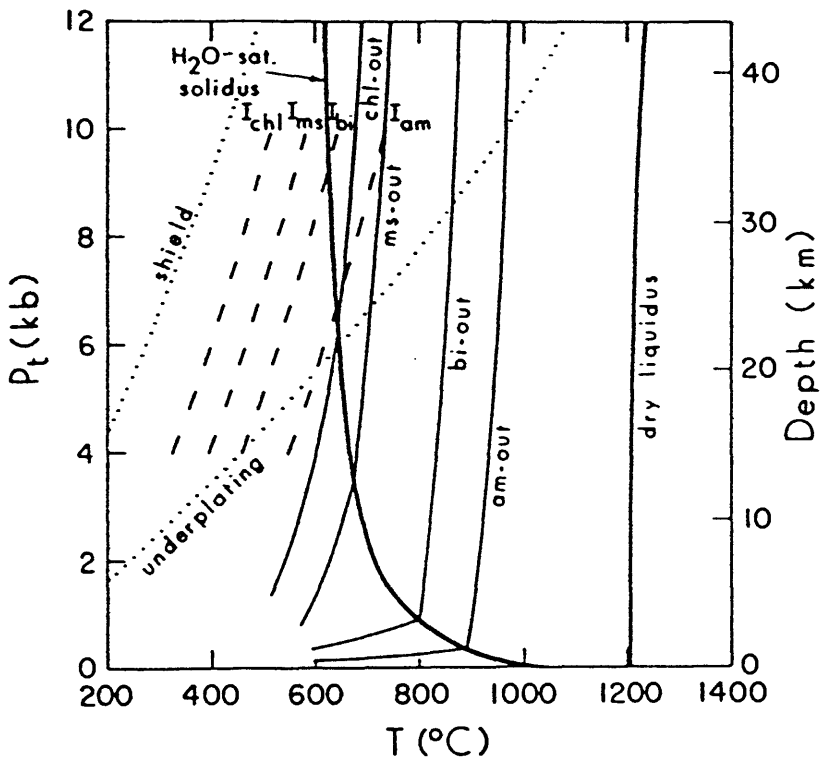



Figure 2. Pressure-temperature projections of melting and fluid production in intermediate composition rocks for various degrees of water undersaturation. Traces of invariant points I (dashed lines) represent first appearance of $H_2O$-$CO_2$ fluid in the presence of hydrate minerals representing various metamorphic grades. Dotted curves are geotherms of shields and magmatically underplated regions (from Wannamaker, 1986).

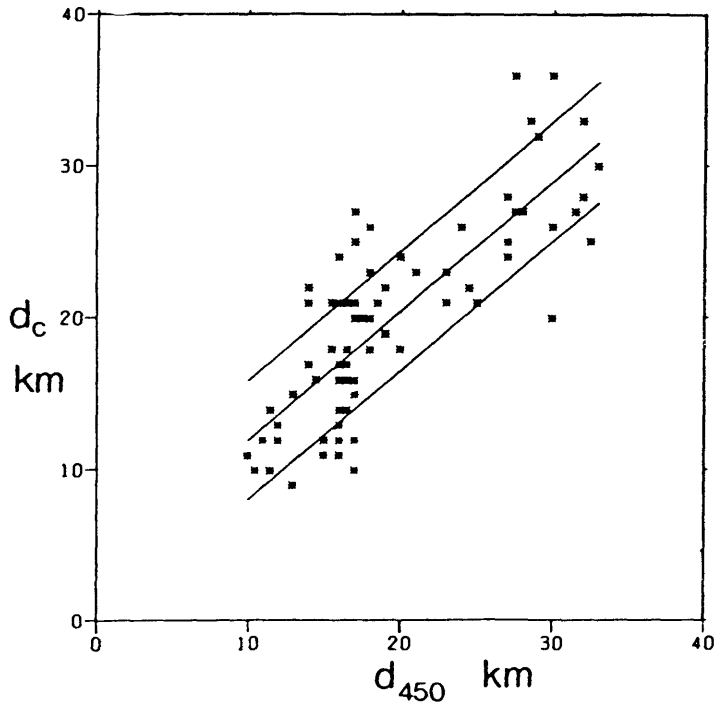



Figure 3. Scatter diagram for the depth $d_C$ to the lower crustal conductor estimated at 76 MT sites in southern British Columbia plotted against depth $d_{450}$ to the 450°C isotherm estimated from heat flow data (from Majorowicz et al., 1993).

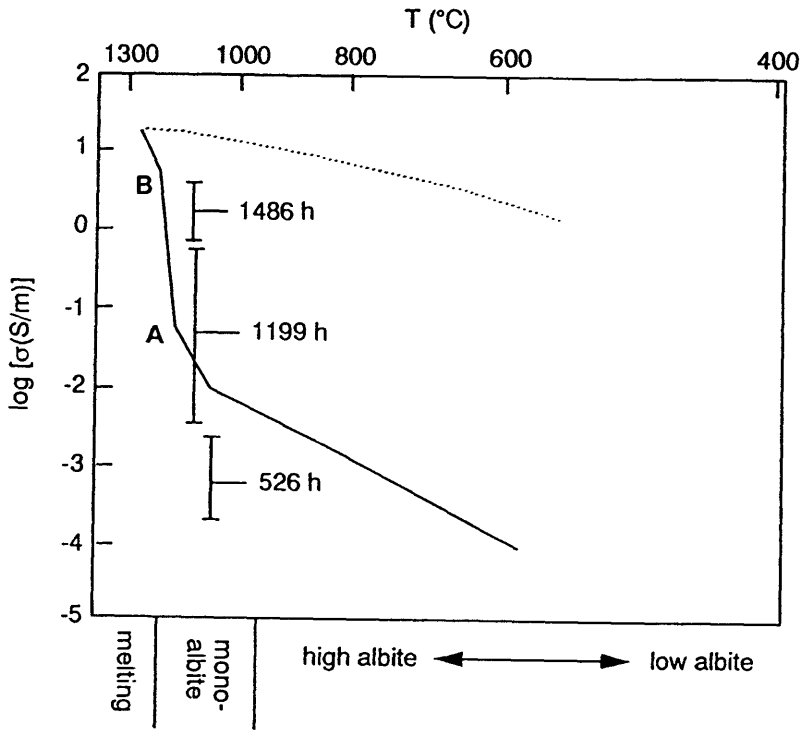



Figure 4. Logarithm of electrical conductivity of polycrystalline albite versus reciprocal temperature for both heating and cooling cycles, and as a function of time in hours at a single temperature (vertical lines). At bottom are labeled solid-state phase relations of albite as a function of defect structure. Modified from Duba (1976).

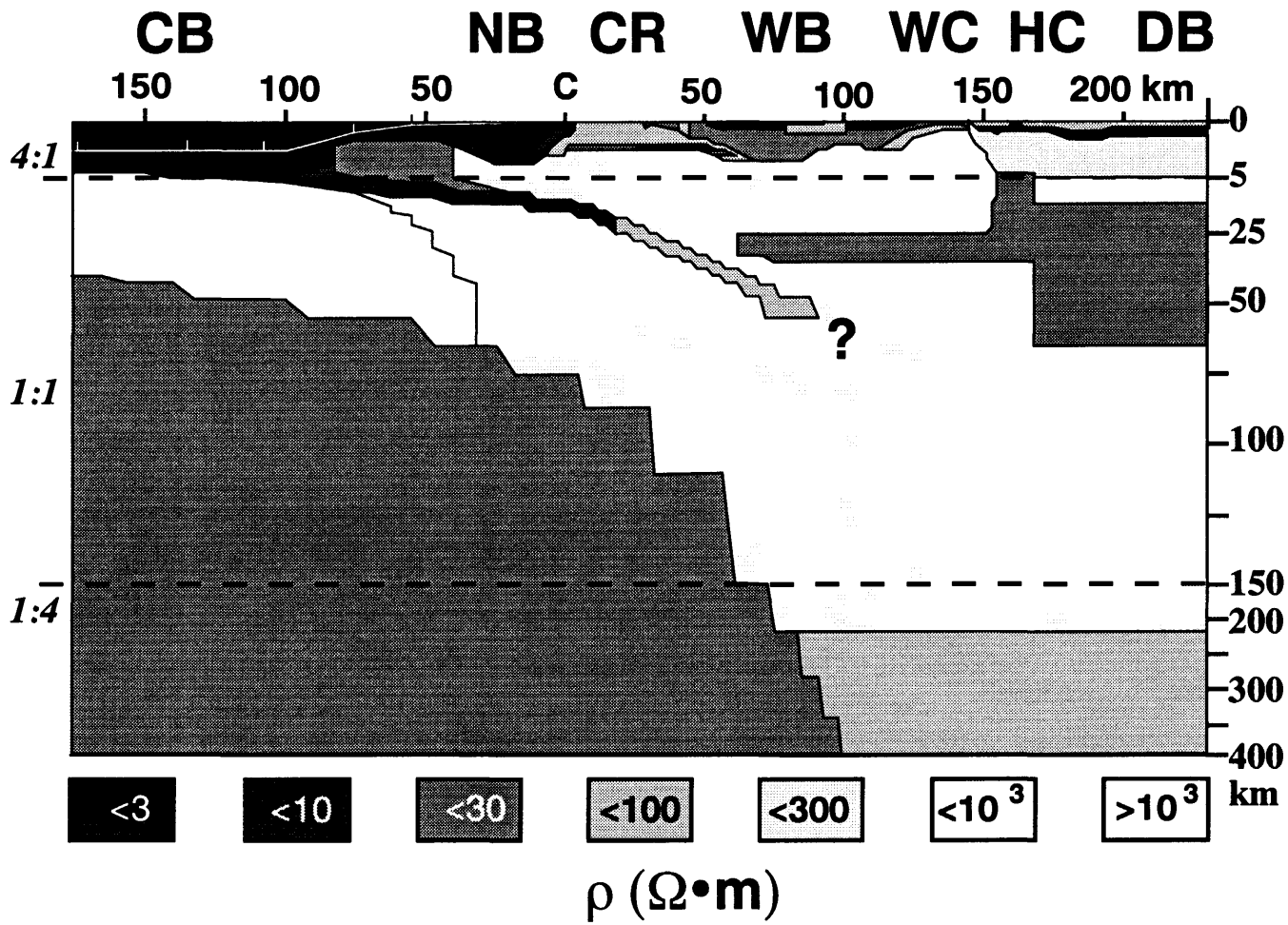



Figure 5. East-west resistivity cross section derived from MT data in the EMSLAB project (modified from Wannamaker et al., 1989). Important physiographic regions crossed include the Cascadia Basin (CB), Newport Basin (NB), Coast Range (CR), Willamette Basin (WB), Western Cascades (WC), High Cascades (HC), and the Deschutes Basin (DB).

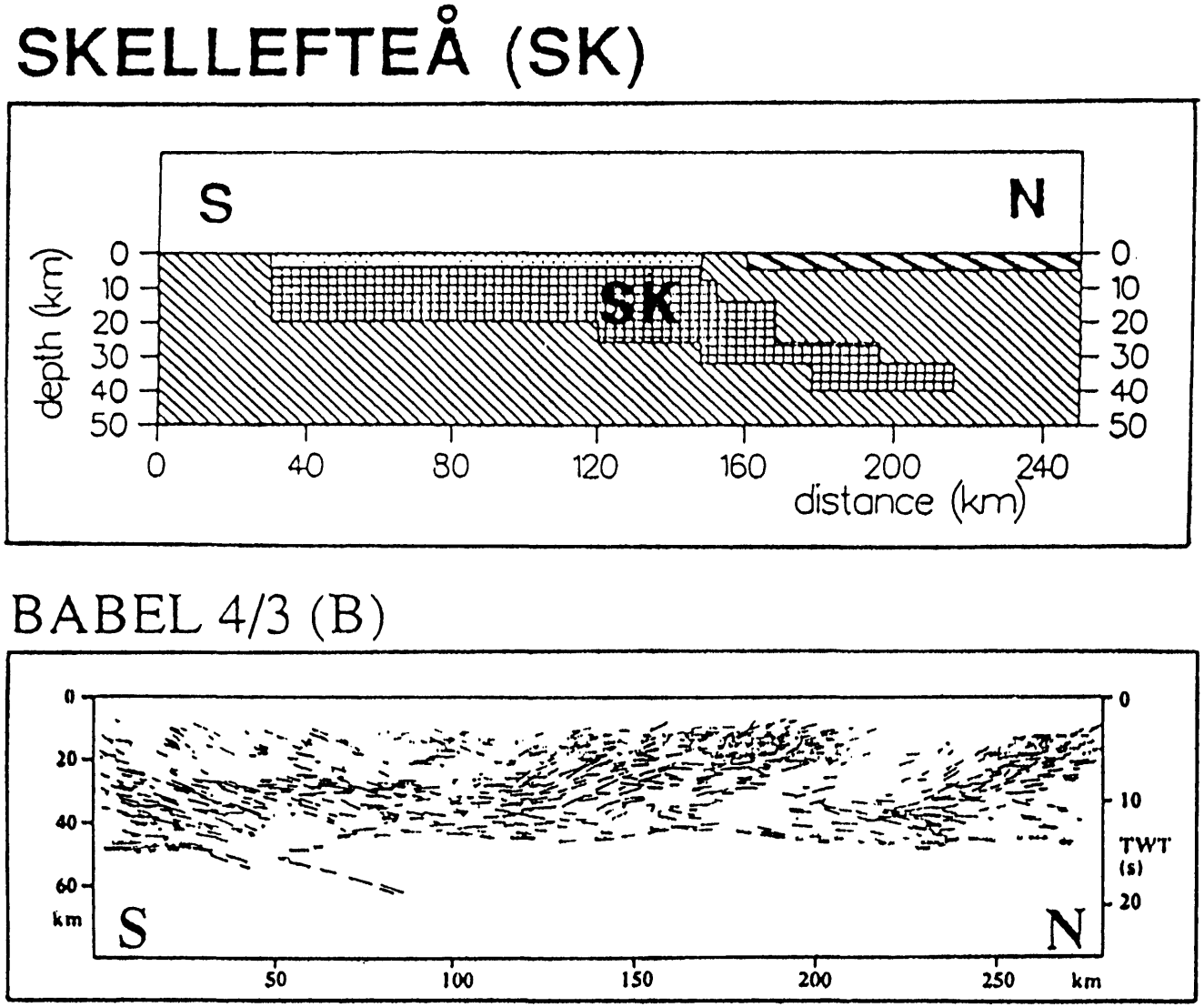



Figure 6. Resistivity cross section through the Svecofennian-Archean suture zone in northern Scandinavia (top) and reflection seismic data from the BABEL project (from Korja et al., 1993). North-directed subduction lead to both underthrust and obducted metasedimentary rocks.

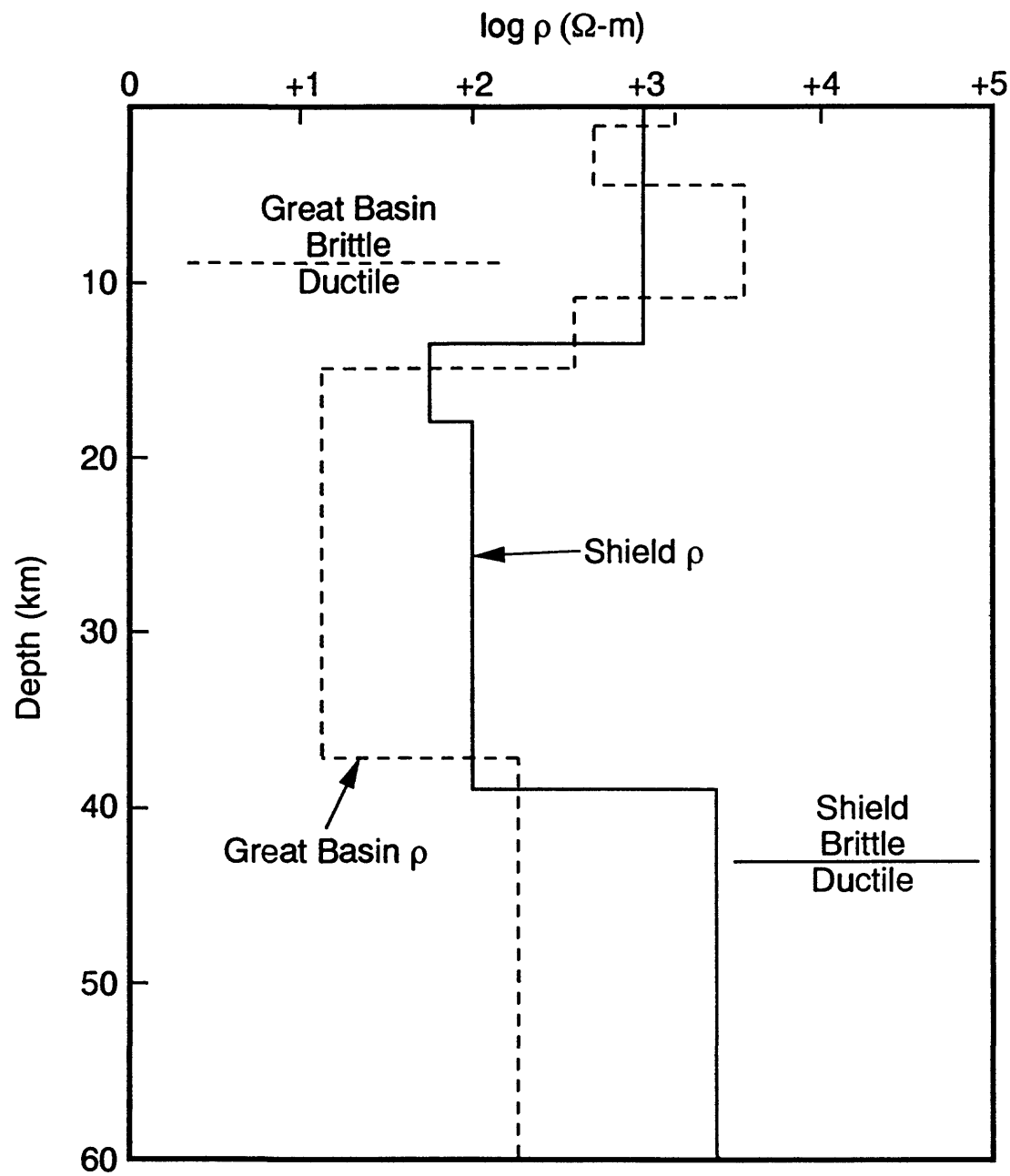



Figure 7. Resistivity profiles through the crust in the Canadian shield of eastern Ontario (solid lines, Kellett et al., 1992) and in the extensional eastern Great Basin of Utah (dashed lines, Wannamaker and Johnston, 1991). The two profiles show low-resistivity layers at similar depths but probably due to different physical conditions.

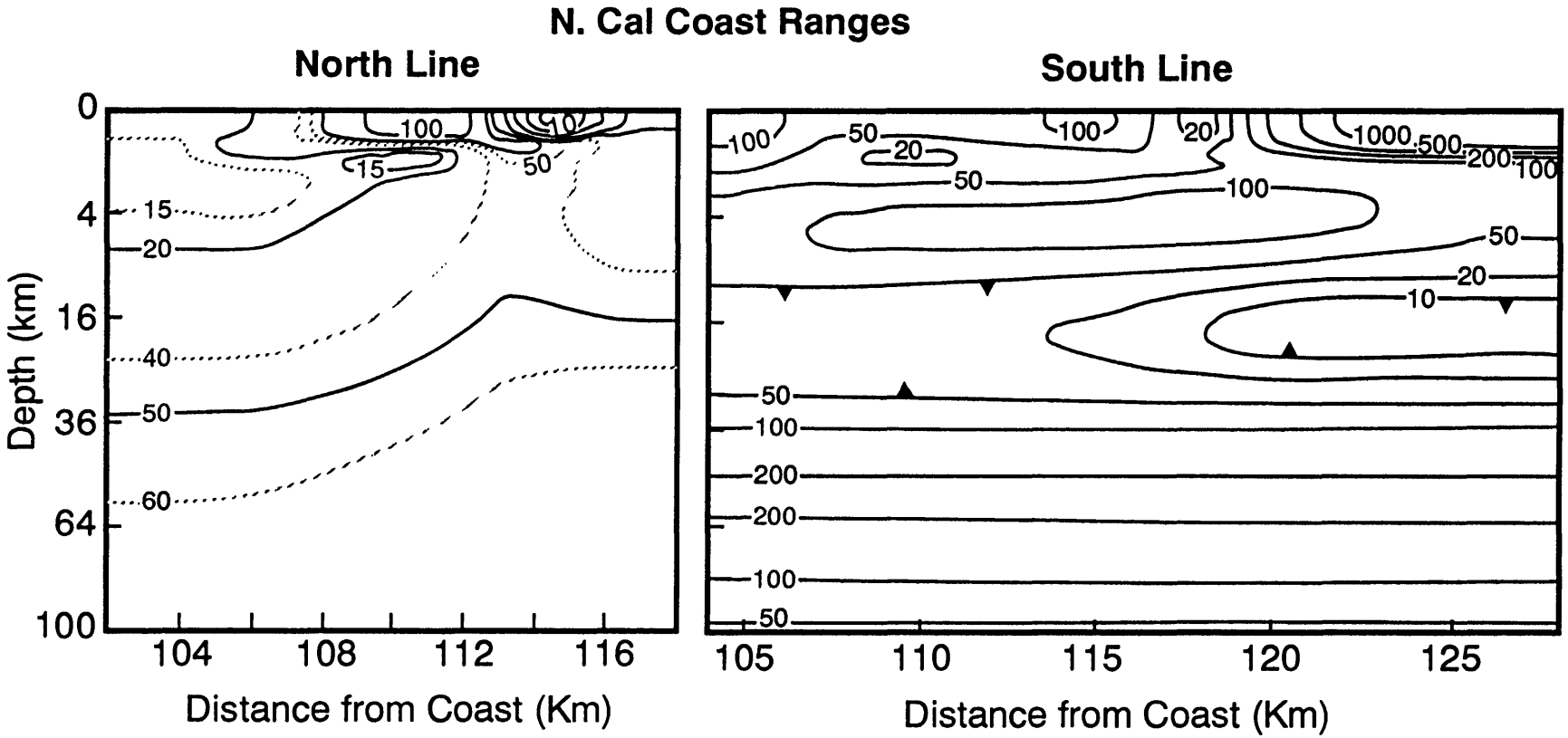





Figure 8. Resistivity models for the crust beneath the northern California Coast Range. The northern and southern profiles lie respectively north and south of the southern edge of the Gorda Plate and represent crust subject to differing thermal inputs in recent geological time (Booker et al., 1992).

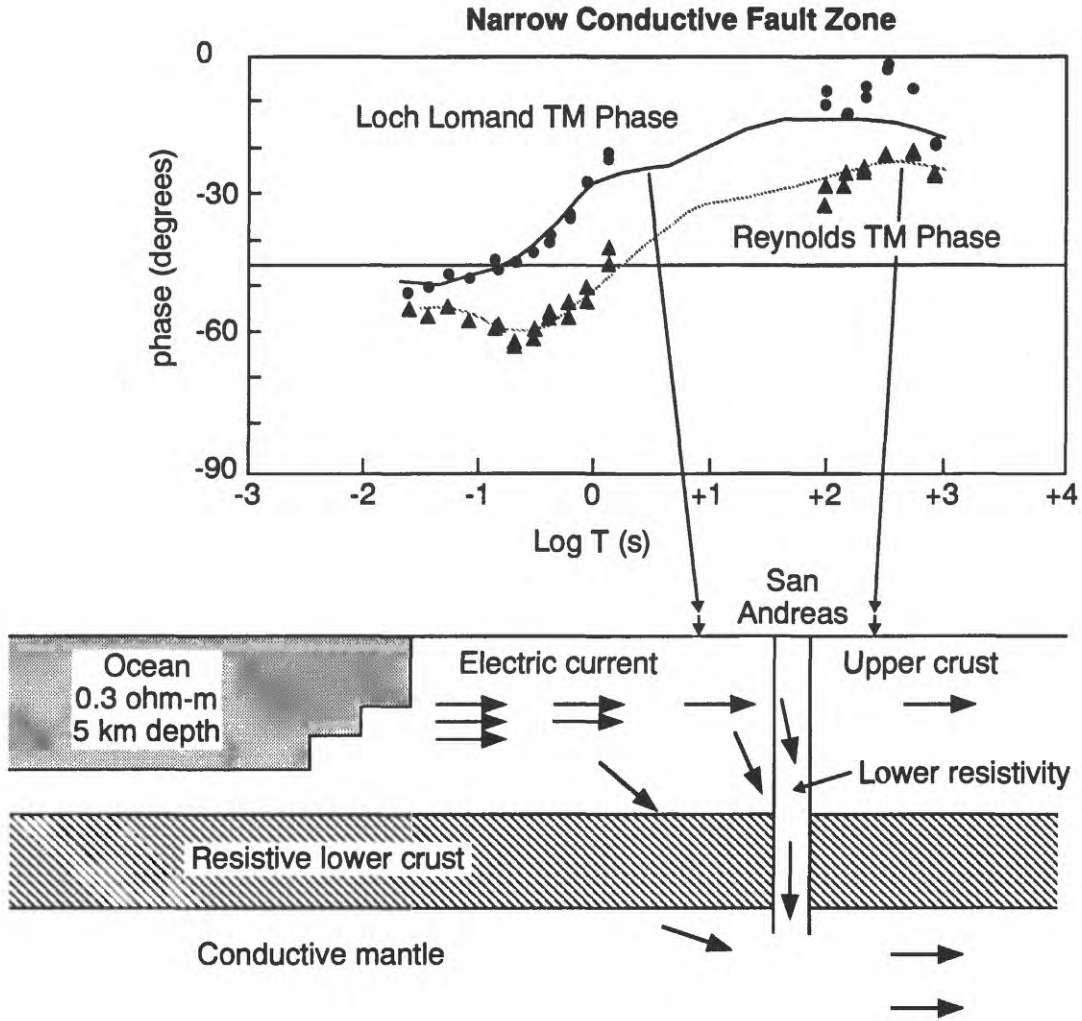



Figure 9. Simplified representation of the crustal resistivity of the central California coastal region including the San Andreas fault zone. Channeling of electrical currents produces nearly discontinuous MT responses for sites straddling the fault zone (redrawn from Mackie et al., 1993).

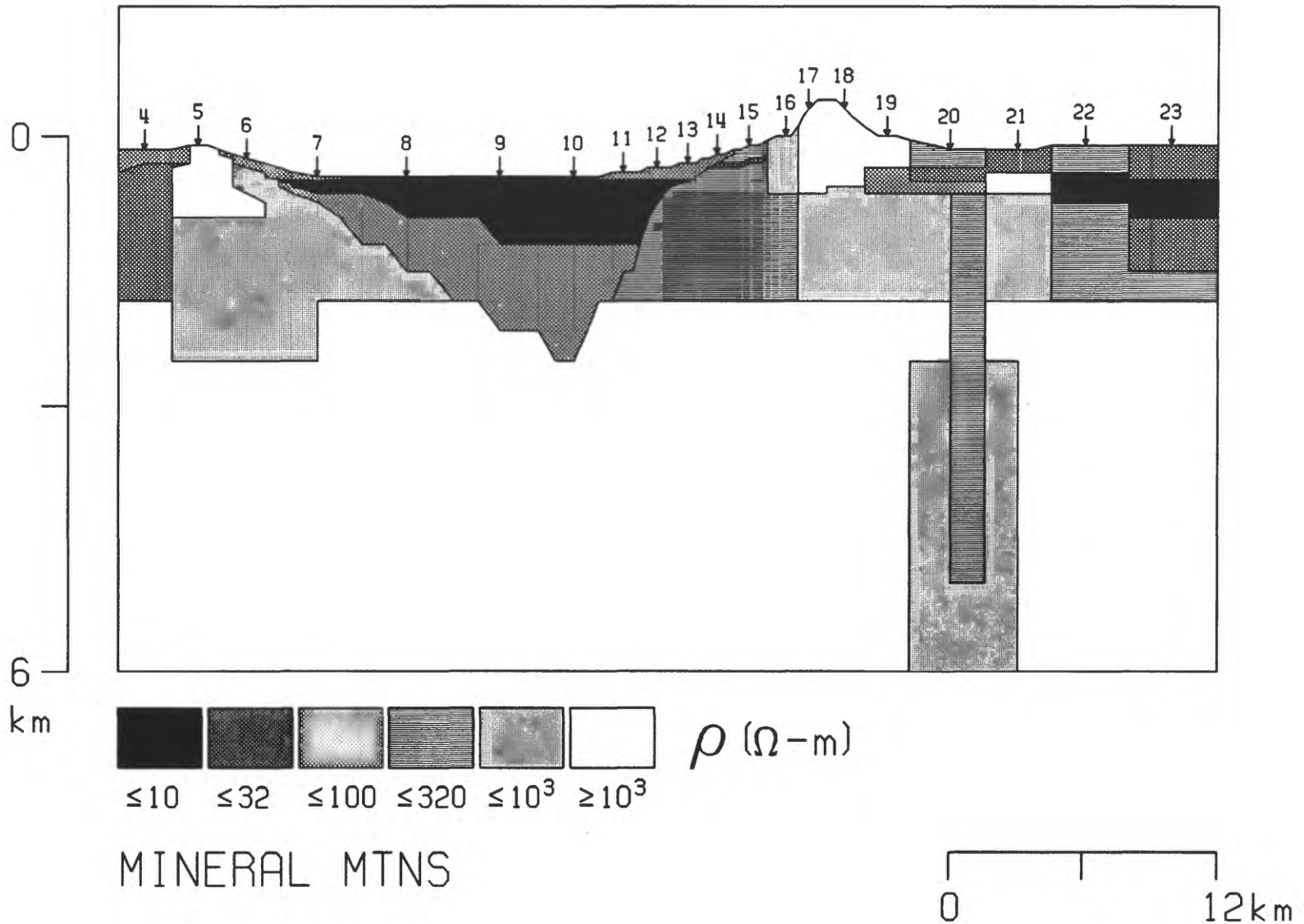



Figure 10. Resistivity model of easternmost twenty sites of the Great Basin study area of Wannamaker and Johnston (1991) covering the Milford Valley and Mineral Mountains.

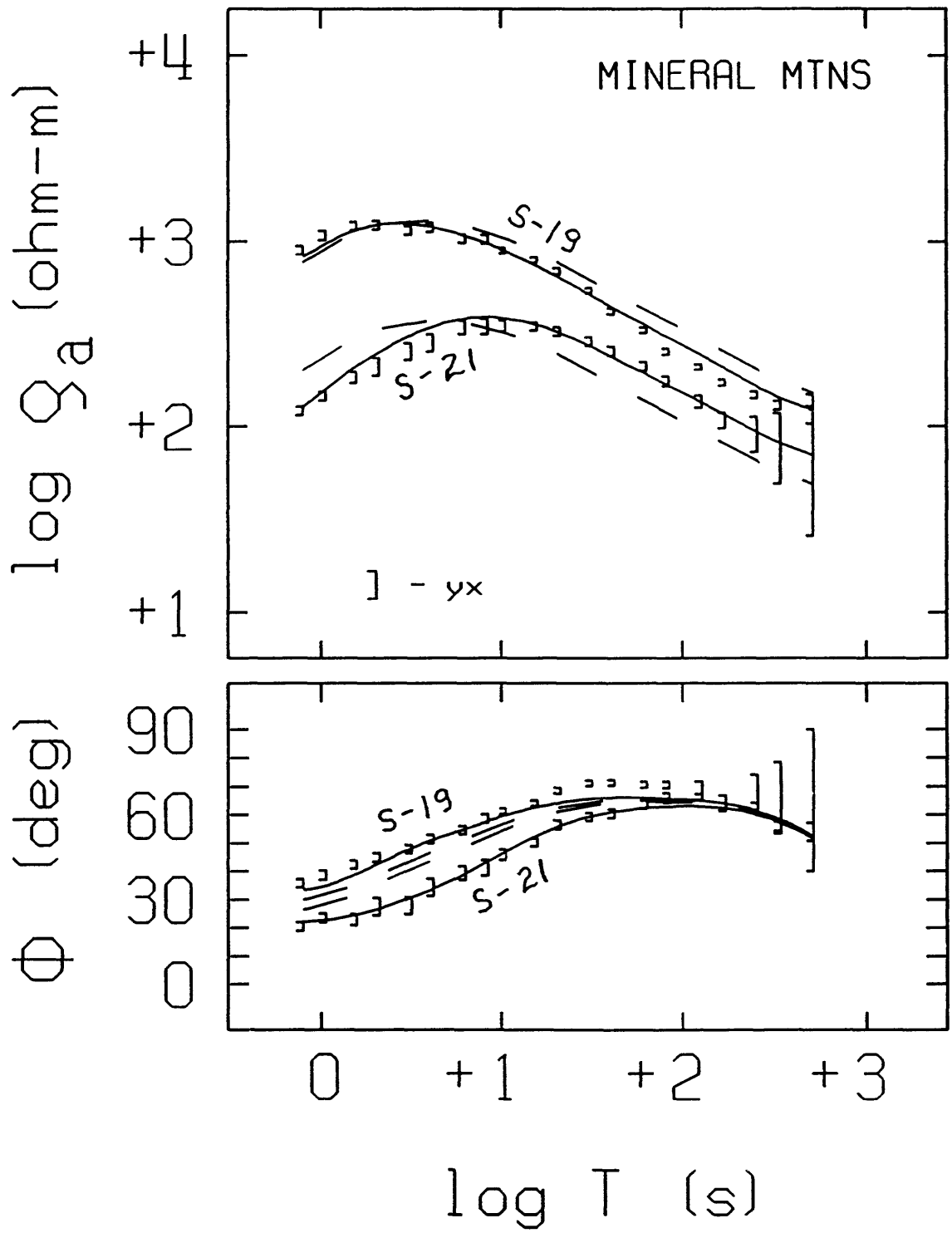



Figure 11. Transverse magnetic (TM) data from soundings S19 and S21 straddling possible projection of Hurricane fault zone into Great Basin study area. Dashed curves indicate poorer fit when such a pathway is removed and best-fitting structure otherwise sought.

# Widespread Fluids in the Lower Crust: A Source to Crustal Penetrating Faults?

**R.D. Hyndman**

*Pacific Geoscience Centre, Geological Survey of Canada*
*P.O. Box 6000, Sidney, B.C., V8L 4B2, Canada*

*Abstract*

One explanation for the very low strength and low effective coefficient of friction inferred for transcurrent and thrust faults is the presence of high fluid pressure. The extensive indirect evidence for a few percent of widespread pervasive water in the lower crust is summarized in this article. The most important evidence is the regional high electrical conductivity of the lower crust as determined by magnetotelluric measurements. However, aspects of seismic reflection and refraction data for the deep crust can also be interpreted in terms of a few percent porosity. The main objection to the presence of such water is the inference that most of the lower crust is made up of dry granulite facies rocks which are electrically resistive. This has led to the main alternative explanation for the high conductivity of thin graphite films. If fluid porosity is present in the lower crust it must be at close to lithostatic pressure and it could provide a deep high fluid pressure reservoir that is tapped by crustal penetrating transcurrent and thrust fault zones.

## INTRODUCTION

Major crustal penetrating faults such as the San Andreas have been shown to be remarkably weak, from the lack of evidence for frictional heating in heat flow data (e.g., Lachenbruch and Sass, 1980; Lachenbruch and Sass, 1992), and from the rotation of the stress field to nearly perpendicular to the fault as the fault trace is approached (e.g., Zoback et al., 1987; Zoback and Beroza, 1993). One explanation for the very strength and low effective coefficient of friction in the fault is the presence of high fluid pressures. It has also been argued that the fluid pressure within the fault zone must be greater than in the surrounding rock (e.g., Rice, 1992), which seems to require an ongoing source of fluid in the deep crust feeding the base of the fault. Here I review the evidence for a few percent of widespread pervasive water in the lower crust. The most important evidence is the generally high electrical conductivity of the lower crust which is readily explained by a few percent interconnected saline fluid porosity (e.g., Shankland and Ander, 1983; Haak and Hutton, 1986; Vanyan and Shilovski, 1989). The main objection to the presence of such water is the inference that most of the lower crust is made up of dry granulite facies rocks (e.g., Yardley, 1981; 1986) which are electrically resistive (Wannamaker, 1986). This has led to the main alternative model of conductive thin graphite films; I will not discuss this alternative here (see Hyndman et al., 1993, for a discussion of the model alternatives).

## HIGH ELECTRICAL CONDUCTIVITY IN THE LOWER CRUST

Magnetotelluric measurements show the lower continental crust to be surprisingly electrically conductive globally, typically 100-1000 times more conductive compared to dry rocks measured in the laboratory at insitu conditions, and compared to the middle to upper crust (e.g., Shankland and Ander, 1983; Haak and Hutton, 1986; Vanyan and Shilovski, 1989; Marquis and Hyndman, 1992). Although there is considerable variability, the mean for regions where the last thermotectonic event was in the Phanerozoic is about 30 Ohm-m, and that for Precambrian areas about 300 Ohm-m (**Figure 1**). Some of the scatter in conductivity is associated with measurement and analysis error. A particular problem is that a thin high conductivity layer and a thicker less conductive layer are not readily distinguished. These conductivities are in contrast to those of dry rocks measured in the laboratory at deep crustal conditions, i.e., about $10^5$ Ohm-m for common rocks expected to occur in the lower crust (e.g., Kariya and Shankland, 1983) (**Figure 1**). However, only a very small amount of interconnected saline porosity, which will have a resistivity of about 0.02 Ohm-m in the lower crust (Nesbitt, 1993), can reduced the bulk rock resistivity to the measured range.

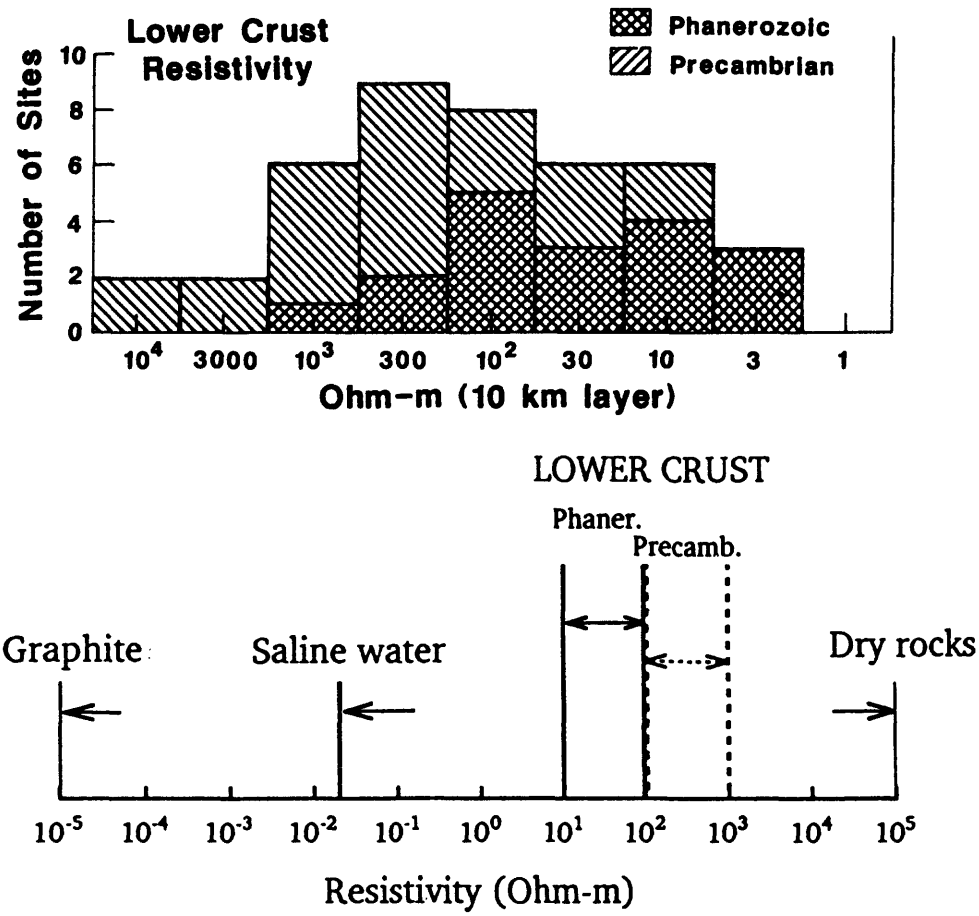



**Figure 1. (a) Histogram of lower crust resistivity data showing the contrast between Phanerozoic and Precambrian regions; (b) Comparison of resistivities in the lower crust to those of graphite, seawater salinity fluid, and common dry rocks.**

Some important characteristic of the conductive lower crust are:

1. The resistivity and depth to the conductive layer show little change across many large scale geological terrane and structural boundaries; this suggests that the conduction is not controlled by the bulk composition.
2. The lower crust on average is more conductive in Phanerozoic areas compared to Precambrian areas (10-30 vs 100-300 Ohm-m, if layer is about 10 km thick, **Figure 1**); Precambrian areas may have experienced multiple thermal dehydration events and a reduction of fluid supply so their lower crusts contain much less fluid and thus are more resistive.
3. The depth to the top of the conductive layer is generally shallower in young areas compared to Precambrian regions (10-20 vs 20-35 km); the depth appears to be correlated to heat flow and commonly corresponds with a present temperature of 350-400°C (**Figure 2**). The conduction mechanism thus must be thermally controlled.
4. In some areas there appears to be correlations between the resistivity of the lower crust and seismic reflectivity, seismic velocity, and seismic attenuation.

All of these characteristics of the deep crustal conductive layers can be explained by a few percent porosity that is restricted to below the brittle-ductile transition at about 400°C. They are not readily explained by the alternative explanation of thin interconnected graphite films (e.g., Hyndman et al., 1993).

## FLUIDS AND SEISMIC CHARACTERISTICS OF THE LOWER CRUST

Seismic data for the lower crust does not provide such clear evidence of fluids in the lower crust as does the results of electrical measurements; however, two seismic characteristics can be interpreted in terms of a few percent porosity. Firstly, the lower crust appears to have too low a seismic velocity (avg. 6.7 km/s) from refraction measurements compared to that for a mafic composition as indicated by lower crust xenoliths ($\sim$7.2 km/s); this discrepancy can be resolved by a few percent porosity (Hyndman and Klemperer, 1989) (1% porosity will reduce velocity by about 10% if equilibrium pores occur). However, the discrepancy between the observed and expected velocities in the lower crust are within the uncertainties, so this evidence is not uncontrovertible.

Secondly, the lower crust, particularly in Phanerozoic areas, is commonly reflective compared to the upper and middle crust; horizontally layered porosity could explain this reflectivity (e.g., Gough, 1986; Hyndman and Shearer, 1989). While layered porosity is not the most commonly accepted explanation for the reflectivity, this mechanism seems to be the only one that explains why the reflective crust most commonly occurs at depths where the present temperature exceeds 350-400°C, a temperature similar to that for the top of the conductive lower crust (**Figure 2**). Explanations involving mafic layering or mylonite zones do not easily accommodate the evidence for a top to the reflective zones that depends on present rather than past temperature.

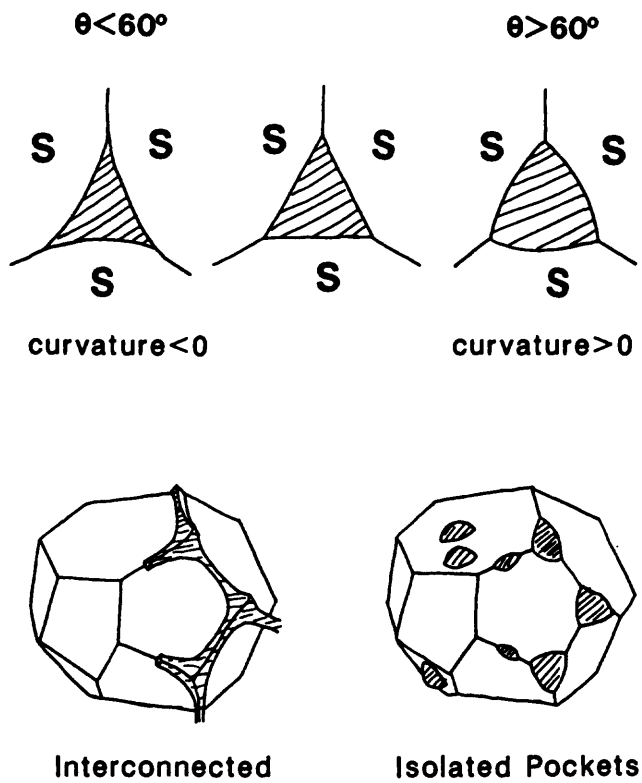



**Figure 2. Textural equilibrium pore geometries expected at temperatures above about 350-400°C. The pores are interconnected for dihedral or wetting angles less than about 60° and pinched off for angles more than 60°.**

## HIGH CONDUCTIVITY AND EQUILIBRIUM PORE GEOMETRIES

High conductivity in the deep crust can be explained by a saline interconnected grain boundary fluid porosity in the range of 0.1 to 2% (e.g., Hyndman and Shearer, 1989). It is suggested that pore fluids are trapped only at depths below the "brittle-ductile" transition, i.e., 350-400°C for probable lower crustal rocks. At such deep crustal temperatures, insitu pore shapes should approach that of textural equilibrium, i.e., minimum energy. Assumption of such porosity allows semi-quantitative estimates of the amount of porosity required to give the observed resistivities. At temperatures greater than about 350-400°C, the grain matrix will have negligible strength over geological times and pore pressures must approach lithostatic. This temperature corresponds well with the temperature estimated at the top of the conductive layers in global compilations (**Figure 3**). An important requirement of this model is horizontal pore interconnection to achieve high electrical conductivity and low vertical pore interconnection (i.e., pinched off) to prevent very rapid upward loss of fluids at lithostatic pressures.

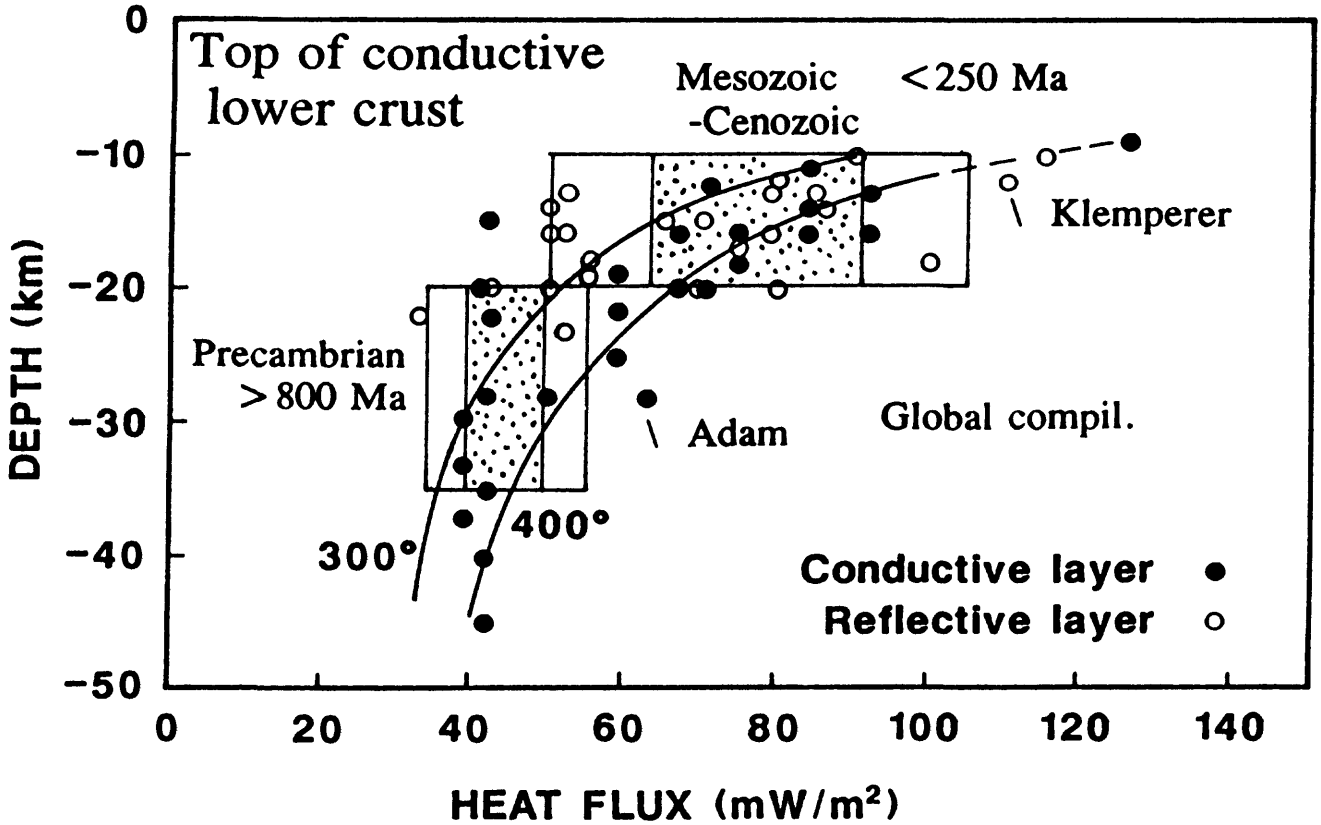



**Figure 3. The depth to the top of electrically conductive and seismically reflective lower crust compared to the local heat flow. The tops are inferred to be at 350-400°C (after Hyndman et al., 1993).**

## SOURCE OF FLUIDS

Three sources of fluids in the lower crust are suggested:

1. Metamorphic fluids; for example dehydration from greenschist to amphibolite facies releases 1 to 2% water. This mechanism requires increasing temperature and has the problem that most regions are at present cooling from past thermotectonic events.
2. Subduction zone fluids, from underthrust sediments, oceanic crustal porosity, and hydrated minerals; the oceanic crust is estimated to contain about 5% water. This mechanism is restricted to subduction zone margins but can explain the high electrical conductivity layers observed beneath the forearc regions of subduction zones (e.g., Hyndman, 1988; Wannamaker et al., 1989).
3. Exsolution from crystallizing magmas or partial melts; melts may contain several percent water in solution. The melts may occur either in the lower crust, for example from mafic underplating, or in the upper mantle, for example from partial melts in a cooling asthenosphere. Mafic underplating has been suggested to be a common phenomenon in extensional environments, especially near the sites of mantle plumes.

## TAPPING LOWER CRUSTAL FLUIDS BY FAULTS

An important problem with the saline fluid model for the electrical conductivity noted above, is that horizontal pore interconnection is required for the high conductivity, and vertical pore pinch-off is required for the fluid not to be lost quickly upward. The requirement can be met by small stresses if the maximum deviatoric stress is horizontal. The surface tension forces that control the textural equilibrium geometry are small and a small horizontal compressive stress of only a few hundred bars will pinch off the vertical pores, leaving the horizontal pores interconnected (Marquis and Hyndman, 1992) (**Figure 4**). For example, the California region of the San Andreas fault has probably been primarily under compression since the change of Pacific-America vector to oblique compression at ~3-5 Ma.

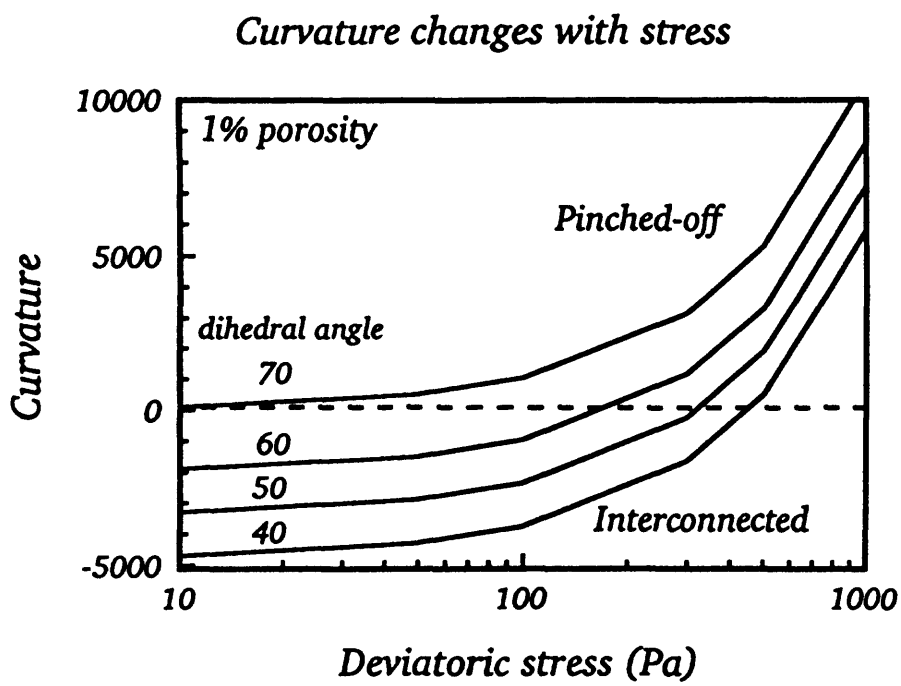



**Figure 4. The change in grain pore curvature as a function of deviatoric stress. For the expected dihedral angle of about 60°, about 300 bars horizontal compression pinches off pores.**

## TRANSCURRENT FAULTS

High fluid pore pressures have been suggested as the origin of fault weakness since the work of Hubbert and Ruby (1959) who mainly studied continental thrust faults. A

mechanism for trapping fluids in the deep crust and releasing them periodically in transcurrent fault earthquakes has been described by several authors (e.g., Rice, 1992; Byerlee, 1990). In these models, the pore pressure is high and approximately equal to the fault normal stress within the fault zone, but decreases with distance into the adjacent crust. If the fault permeability is greater than in the adjacent rock (**Figure 5**), and fault permeability decreases with increasing normal stress, surges of pore pressure may propagate upward along the fault. The source of the fluids may be in the ductile roots of the faults zones. One possibility is that earthquakes temporarily relieve compressive stresses and allow upward migration of pulses of fluids from pervasive lower crustal porosity into fault zone channels.

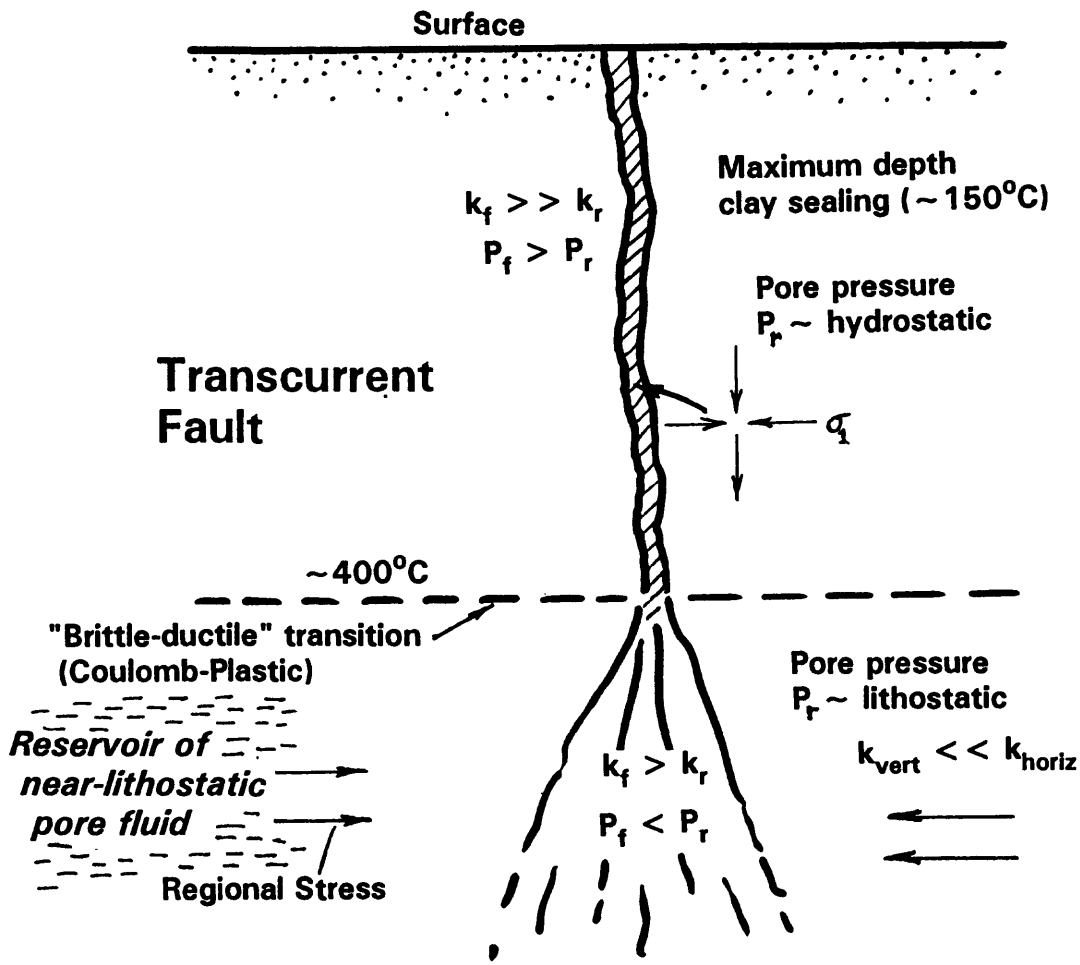



**Figure 5. A possible hydraulic regime for transcurrent faults. In the fault root, the permeability of the fault is higher, but the pore pressure is lower, than in the surroundings. In the brittle regime of the upper crust, the fault zone has higher permeability than the surroundings. However, the pore pressure is also higher, maintained by near-lithostatic fluid supply from below.**

## SUBDUCTION THRUST FAULTS

Although they have not been studied to the same detail as transcurrent faults such as the San Andreas, and continental thrust faults, subduction thrust faults also appear to be weak (e.g., Byrne and Fisher, 1990). A low effective coefficient of friction is also indicated by the apparently negligible amount of frictional heating that they generate. As an example, the very low heat flow in the forearc region of the Cascadia subduction zone ($<35$ mW/m$^2$) can only be modelled with very little frictional heating (Hyndman and Wang, 1993). Low strength is also argued to explain the presence of brittle deformation in subduction zone accretionary prisms (e.g., Karig, 1990). With increasing depth, the principle stress is inferred to rotate from approximately horizontal in the upper portion of a prism (from convergent tectonic stress) to nearly vertical adjacent to the weak decollement fault (**Figure 6**). A similar process to that described for transcurrent fault zone may explain their apparent weakness, i.e., high pore pressures that originate at great depth.

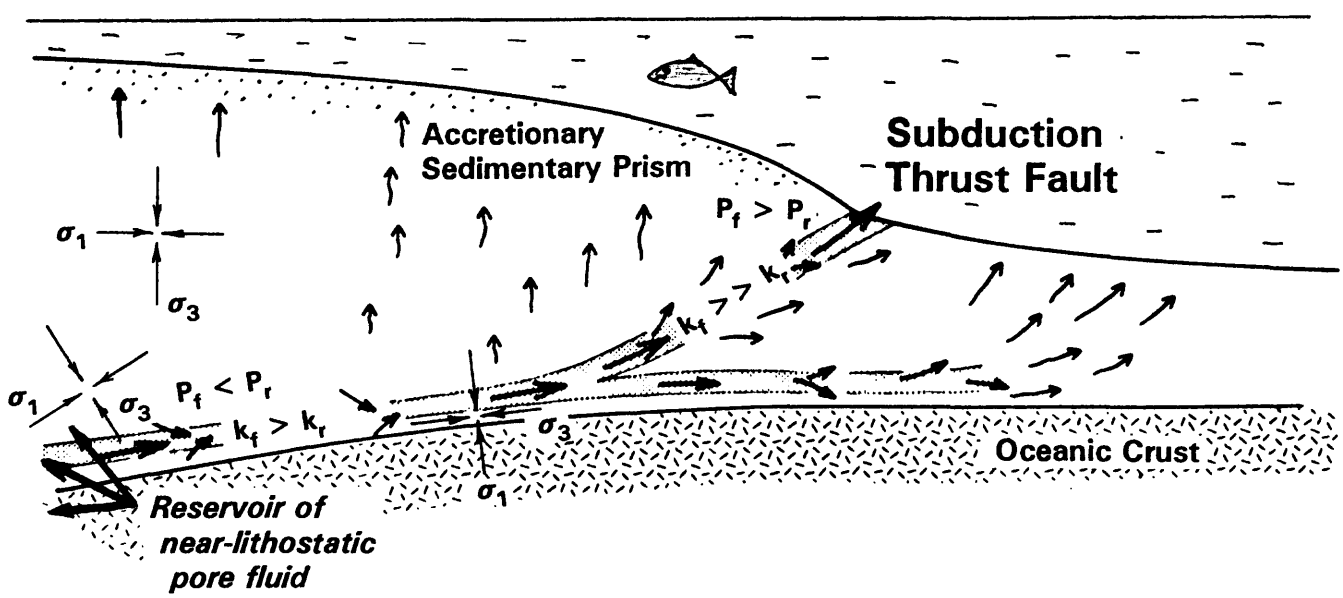



**Figure 6. Subduction thrust faults Modified from Karig (1992). With increasing depth, the principle stress is inferred to rotate from approximately horizontal in the upper portion of an accretionary sedimentary prism (from convergent tectonic stress) to nearly vertical adjacent to and below the weak decollement fault**

## CONCLUSIONS

The low strength and low effective coefficient of friction of continental transcurrent and thrust faults may result from high pore pressure fluids carried up from the lower crust. Evidence for a few percent pore fluid in the lower crust (~0.1-0.5% in Precambrian, ~0.5-3% in Phanerozoic areas, perhaps over a 10 km thick layer) is provided by the high electrical conductivity that is determined by magnetotelluric measurements in many regions. Several aspects of seismic reflection and refraction data for the deep crust can also be interpreted in terms of a few percent porosity. An important conclusion from the correlation of the depth to the top of the conductive deep crust (and deep seismic reflectivity) with surface heat flow, is that the layer occurs at temperatures above about 400°C, i.e., deeper in low heat flow areas, shallow in high heat flow areas. If the conductivity and reflectivity are produced by horizontally interconnected porosity, such porosity only occurs in the high temperature ductile regime. Thus, if fluid porosity is present in the lower crust, it must be at close to lithostatic pressure and it could provide a deep high fluid pressure reservoir that is tapped by crustal penetrating transcurrent and thrust fault zones.

## REFERENCES


Byerlee, J.D., 1990. Friction, overpressure and fault normal compression, Geophys. Res. Lett., 17, 2109-2112.

Byrne, T., and D. Fisher, 1990. Evidence for a weak and overpressured decollement beneath sediment-dominated accretionary prisms, J. Geophys. Res., 95, 9081-9097.

Gough, D.I., 1986. Seismic reflectors, conductivity, water and stress in the continental crust, Nature, 323, 143-144.

Haak, V., and R. Hutton, 1986. Electrical resistivity in continental lower crust, in Nature of the Lower Continental Crust, ed. Dawson, J.B., D.A. Carswell, J. Hall, and K.H. Wedepohl, Spec. Publ. 24, Geol. Soc. Lond., 35-49.

Hubbert, M.K., and W.W., Rubey, 1959. Role of fluid pressure in the mechanics of overthrust faulting I: Mechanics of fluid-filled porous solids and its application to overthrust faulting, Geol. Soc. Am. Bull., 70, 115-166.

Hyndman, R.D., 1988. Dipping seismic reflectors, electrically conductive zones and trapped water in the crust over a subducting plate, J. Geophys. Res., 93, 13391-13405.

Hyndman, R.D., L.L. Vanyan, G. Marquis, and L.K. Law, 1993. The origin of electrically conductive lower continental crust: water or graphite? Phys. Earth Planet. Int., in press.

Hyndman, R.D., and P.M. Shearer, 1989. Water the lower continental crust: modelling magnetotelluric and seismic reflection results, Geophys. J., 98, 343-365.

Hyndman, R.D., and S.L. Klemperer, 1989. Lower crustal porosity from electrical measurements and inferences about composition from seismic velocities, Geophys. Res. Lett., 16, 255-258.

Hyndman, R.D., and K. Wang, 1993. Thermal constraints on the zone of major thrust earthquake failure: The Cascadia subduction zone, J. Geophys. Res., 98, 2039-2060.

Karig, D.E., 1990. Experimental and observational constraints on the mechanical behaviour in the toes of accretionary prisms, in Deformation Mechanisms, Rheology and Tectonics, eds. Knipe, R.J. and E.H. Rutter, Geol. Soc. Spec. Publ. 54, 383-398.

Kariya, K.A., and T.J. Shankland, 1983. Electrical conductivity of dry lower crustal rocks, Geophysics, 48, 52-61.

Lachenbruch, A.H., and J.H. Sass, 1980. Heat flow and energetics of the San Andreas fault zone, J. Geophys. Res., 85, 6185-6223.

Lachenbruch, A.H., and J.H. Sass, 1992. Heat flow from Cajon Pass, fault strength, and tectonic implications, J. Geophys. Res., 97, 4995-5015.

Marquis, G., and R.D. Hyndman, 1992, Geophysical support for aqueous fluids in the deep crust: seismic and electrical relationships, Geophys. J., 110, 91-105.

Nesbitt, B.E., 1993. Electrical resistivities of crustal fluids, J. Geophys. Res., 98, 4301-4310.

Rice, J.R., 1992. Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault, in Fault Mechanics and Transport Properties of Rocks, ed. Evans, B. and T.-F. Wong, Academic Press.

Shankland, T.J., and M.E. Ander, 1983. Electrical conductivity, temperatures, and fluids in the lower crust, J. Geophys. Res., 88, 9475-9484.

Vanyan, L.L., and A.P. Shilovski, 1989. Fluids in the lower crust inferred from electromagnetic data, in Properties and Processes of the Earth's Lower Crust, eds.

Mereu, R., S. Mueller, and D. Fountain, Am. Geophys. Un. Monograph 51, Washington, D.C.

Wannamaker, P.E., 1986. Electrical conductivity of water-undersaturated crustal melting, J. Geophys. Res., 91, 6321-6327.

Wannamaker, P.E., J.R. Booker, A.G. Jones, A.D. Chave, J.H. Filloux, H.S. Waff, and L.K. Law, 1989. Resistivity cross-section through the Juan de Fuca subduction system and its tectonic implications, J. Geophys. Res., 94, 14,127-14,144.

Yardley, B.W., 1981. Effect of cooling on the water content and mechanical behaviour of metamorphosed rocks, Geology, 9, 405-408.

Yardley, B.W., 1986. Is there water in the deep continental crust? Nature, 323, 111.

Zoback, M.D., et al., 1987, New evidence on the state of stress of the San Andreas fault system, Science, 238, 1105-1111.

Zoback, M.D., and G.C. Beroza, 1993. Evidence for near-frictionless faulting in the 1989 (M 6.9) Loma Prieta, California, earthquake and its aftershocks, Geology, 21, 181-185.

# 2: FAULT-ZONE TRANSPORT PROPERTIES AND COMPOSITION OF FAULT-ZONE FLUIDS

# CYCLIC FLUID FLOW ALONG FAULTS

John M. Logan and Carrie L. Decker
Center for Tectonophysics, Texas A&M University, College Station, TX 77843

## Abstract

Field observations of fault and fracture mineralization are presented documenting faults as conduits and barriers to fluid flow. Imposed upon channeling of fluids in the fault gouge, some faults show anisotropic outward flow, with only one side of the fault being mineralized. Cyclic fault displacement, fluid flow, mineralization which produces healing of the gouge, and a later stage of displacement is documented in thin-section observations. Multiple cycles have been recognized. Fluid-inclusion studies support the petrographic evidence of multiple episodes of fluid migration with changing salinities and temperatures of homogenization. These observations suggest that the mechanical properties of the faults were continuously changing during their displacement history, seldom achieving "steady state" behavior.

## Introduction

Fault zones have been invoked as conduits for fluid migration, as barriers to fluid flow, and at times both roles have been ascribed to the same fault. Often one condition has been postulated to be followed by the other, and cyclic situations have frequently been suggested. In brief, a wide range of behavior is envisioned to rationalize field observations or theoretical models. In the upper crust, evidence for fault-zones acting as channels of fluid migration comes from the mining industry, while documentation of faults as seals to fluid migration comes primarily from studies of hydrocarbon reservoirs. Despite this evidence, our understanding of the role of fault zones in fluid flow remains meager. An even greater paucity of information exits on the affect of fluids on the mechanical behavior, for instance the role of fluid chemistry on strength and stability.

The latter have been explored in an earlier work (Feucht and Logan, 1990). The former issue--faults as conduits and barriers to fluid flow--is addressed here through recent field and laboratory studies. Specifically, evidence for cyclic fault movements alternating with separate periods of fluid migration and mineralization is presented.

## Fault Zone Geometry and Fabric

Previous studies have established a cross-sectional geometry that has been found to typify many faults within the upper portion of the crust (Figure 1; Wallace and Morris, 1986; Logan, 1992 ). The nominal fault zone width, or that which is often mapped as the "fault" on the ground, is a zone of fracturing where the fracture density is significantly higher than the country rock on both sides. Within this zone of high fracture density is the fault gouge, a zone of fine-grained material which has been formed by comminution of rock from both sides of the fault. It is within the gouge that the displacement of the fault is accommodated; this is the "active trace". Laboratory studies on simulated gouge combined with field measurements have established a recognizable fabric which is found in many faults (Figure 2; Byerlee et al., 1978; Logan et. al. 1979, 1992; Rutter et. al., 1986; Chester and Logan, 1987)

Within the fault gouge, the grain size is significantly smaller than in the fractured halo or country rock because it is formed primarily by cataclasis where microfracturing is accompanied by rotation and translation of the grains. The porosity and permeability are lower than in the fracture halo immediately adjacent to the gouge. The fractured halo will in turn have a higher porosity and permeability than the country rock, because of the increased fracture porosity and connected pore space. Not only is the general porosity and permeability of the gouge reduced from the country rock, but within the gouge the degree of grain comminution, and thus porosity and permeability are heterogeneous (Figure 3).

This heterogeneity is documented in experimental studies on simulated gouge of quartz which show that the average grain size is reduced by an order of magnitude after an initial shear strain of about 20%. But along the early formed $R_1$, $R_2$ and P fractures they are reduced by

another factor of ten, and along the Y-fractures a third order of magnitude (Shimamoto, 1977). The gouge contains, therefore, domains of higher porosity bounded by fractures where permeability and porosity are much lower suggesting that fluid flow may be heterogeneous and anisotropic within the gouge and the fault in general (Figure 3).

To indicate the potential magnitude of permeability reduction caused by even a small amount of shear strain, measurements on quartz gouge with a starting grain size of about 1 micron and sheared about 5 mm at an effective confining pressure of 150 MPa, resulted in permeabilities in the microdarcy range as measured across the zone (Higgs, 1981). In another set of experiments cylinders of Berea sandstone containing a saw cut parallel to the loading and fluid-flow directions, were taken to failure in triaxial compression; a fault formed intersecting the saw cut. The fault zone was about one millimeter thick and underwent an equal amount of displacement. Air permeabilities decreased from about 850 md to about 250 md with this very small deformation (Gatto, 1984).

These field and laboratory studies show that fault zones have the potential to both restrict and enhance fluid flow. Parallel to a fault, enhancement of flow may occur. This is particularly true where the fracture halo is much wider than the gouge region. Even within the gouge, flow parallel to the fault may be enhanced as Y-fractures parallel the overall strike and dip. Their large aerial extent and anastomosing character over large distances make them ideal fluid conduits. Although these are sites of extreme gain comminution, inhibiting flow normal to them, they are discrete fracture surfaces providing enhanced fluid channels. Transverse to the fault, restriction in flow is predicted as a direct result of the decrease in grain size with a corresponding increase in the tortuosity of the migration path, and the lack of major fractures crossing the gouge once Y-fractures are well developed. Significantly higher values of pore pressure would be expected in the fault gouge than in other portions of the fault zone, or the surrounding country rock. This model can be utilized to aid in interpreting field exposures.

**Faults as Conduits for Fluids and Anisotropic Flow Into the Fracture Halo**

Fault zones in the cliffs and wave-cut terraces in coastal Somerset, UK. provide excellent three dimensional exposures of the relationship between faults and fluids (Figures 4a and b). East-west trending normal faults with tens of meters of displacement cut Jurassic limestones and shales. Over 20 faults have been examined to date. The gouges of most of the faults are mineralized by calcite; some are on the order of 10-20 cm in width while others are less than 5 cm wide (Figures 5 and 6). It is assumed that the presence of calcite is the result of fluids moving through the fault gouge and depositing the mineral. The absence of mineralization does not preclude the presence of fluids at one time, but the probability is presumed to be lower.

The number of mineralized fractures adjacent to the fault gouge decreases rapidly away from a given fault so that the mineralized halo extends only a few meters. It is clear from the concentration of mineralization within the faults and its decreasing abundance normal to them that the faults were formed first and they served to channel calcite-bearing fluids which came from deeper depths. Additional evidence for the latter interpretation is presented below. The fact that mineralization is confined to small halos adjacent to the fault gouge, suggests that fractures were preferentially dilated prior to or during displacement, facilitating fluid migration. This hypothesis is supported by brecciated zones and unfilled vugs adjacent to fault gouge in a number of faults (Figure 6). There is no compelling evidence to support abnormally high pore pressures nor are there field relations that would be better explained by such high pressures.

Superimposed upon the channeling is evidence that the fluid migration from some faults into the adjacent fractures was anisotropic. While often fractures on both sides of the fault are mineralized, frequently mineralization is confined one side (Figures 4a and 7), but there is no evidence that either the hanging or foot wall is preferentially mineralized. This selective mineralization would be enhanced if one side of the fault was dilated more than the other. In addition, the fabric of the fault gouge could produce differential fluid flow out of it. The salient fabric features are Y-fractures, with attending comminuted grains at the edge of the gouge. While these are low-permeability zones, they are not uniform in their properties, that is, restriction of flow will vary from location to location. Y-fractures are found on the edge of the gouge away

from the mineralized fractures. It is inferred that these permeability barriers restricted flow into that wall of the fault but did not on the other.

## Cyclic Fault Displacement and Fluid Migration

Two lines of evidence indicate that fluid migration in the gouge and subsequent mineralization alternate with periods of slip and rupture. The first comes from petrographic observations.

The geologic history of the region suggests extensional tectonics during the formation of the adjacent Bristol channel, but with possible inversion later in its history. The faults then are expected to show normal dip-slip displacement, but some reverse movement could have taken place, overprinting the early offset. Reverse drag is found along some of the faults, but all show dominantly dip-slip motion. The fault shown in Figure 5 is well exposed in both dip and strike section. The latter has a large continuous upper surface of mineralized fault gouge where the calcite contains extensive slickensides (Figure 8). Offset of beds along the fault shows net normal displacement of a few tens of meters. It is clear that the fault was formed creating the generalized fabric within the gouge (Figure 2), and was eventually mineralized by fluids moving up through the gouge. A sample taken from the mineralized gouge was thin-sectioned from the upper slickensided surface into the center of the gouge (Figure 9). In the hand specimen and outcrop, offset markers suggest that displacement was localized along parallel Y-fractures two or three centimeters from the upper slickensided surface. A petrographic traverse from this upper slickensided surface toward the Y-fractures in the interior of the gouge is instructive.

The slickensided layer is fine-grained calcite, with the edges of the slickensides well defined (Figure 10a). Interior to this, is a region of coarser grains, followed by another slickensided-layer. Moving towards the center of the zone, the calcite is still coarse grained with a large number of closely space twins. This high density suggests relatively high differential stresses existed after their precipitation. Additionally, the twins are bent and show undulose extinction indicating a superimposed period of crystal plastic strain. This deformation is overprinted by fractures (Figure 10b). While the twins could form during a period of strain build-up with little displacement of the fault, the subsequent fracturing suggest finite displacement of the fault.

Traversing further into the center of the sample (Figure 10c), the calcite becomes highly fractured with fragments separated. A gradual transition occurs from intact grain boundaries to ones separated and rotated grains. Kerogen is present along separated grain boundaries (Figure 10c). There are oriented fractures inclined to the overall displacement direction of the fault which can be interpreted as P-fractures created during normal dip-slip displacement or $R_1$- fractures of a later reverse displacement phase ( Figure 2). This reflects a second stage of brittle deformation following the formation of the fault.

Further into the fault zone, fine-grained calcite is laminated with kerogen. The calcite veins are parallel to the Y-orientation (Figure 10c) and are interpreted as calcite-filled Y-fractures. The grain size is distinctively finer than the initial mineralization, with few twins and with an oblique preferred orientation; they clearly represent a second cycle of fluid migration. It is inferred that mineralization occurred along Y-fractures formed during a third stage of brittle deformation. One can see that this site which is where most of the fault displacement took place is only a few millimeters wide. Traversing further, clasts of calcite are rotated and elongated in the P-orientation, assuming normal dip-slip displacement (Figure 10d). This feature is very similar to that produced in laboratory experiments on gouge from the San Andreas fault (Logan et al., 1981). Finally, veins cross-cutting all other features at high angles are present. These filled fractures appear to be extensional features formed during strain or displacement of the fault, with a third generation of fluids filling them. The calcite mineralization found in the fault gouge is distinctly different from that found as simple vein fillings. Although the veins are associated with faults they are uniform in their mineralization with only one generation generally present.

In summary the cyclic nature of the deformation and fluid migration with crystallization is inferred to be:

- A first, brittle deformation event forming the fault and fracture halo
- Fluids moving up the fault gouge depositing coarse-grained calcite
- A period of relatively high differential stress, during which closely spaced twins formed
- Straining of the fault, bending the calcite twins
- Brittle deformation, the second event, during which calcite grains were fractured and rotated, and P or $R_1$ fractures formed.
- A migration of kerogen up the fault separating the fractured grains and filling fractures
- A third phase of brittle deformation, during which Y-fractures were formed, and most of the displacement appears to have occurred.
- A second influx of calcite-bearing fluids up the fault precipitating fine-grained crystals along Y-fractures; some syn-deformation apparently accompanying this with rotation of these calcite veins to P-orientations; possibly slickensides formed at this time
- A fourth brittle fracture phase during which the cross-cutting extensional fractures formed
- Third and final calcite mineralization to fill the cross-cutting veins

In addition to this petrographic evidence, fluid-inclusion analysis of this fault and one other show two to three populations of inclusions in each fault. This separation is based on cross-cutting relationships among the inclusions (Decker, et. al., 1992). Each population has a distinct composition and temperature of homogenization (Figure 11). All inclusions are highly saline with 10 to 20 weight percent equivalent NaCl or 3-7 times the salinity of sea water. The trend is for salinity to decrease with time, or slip on the fault, while the homogenization temperature increases. The latter reflects either an increase in fluid temperature or a decrease in pressure. In one of the faults, inclusions of sphalerite are also found, indicating elevated temperatures (about 200°C). This temperature is in contrast to maximum temperatures reached by the bulk rock, which are probably on the order of 60°C, based on a vitrinite reflectance of Ro = 0.4%. This supports the previous inference that the fluid migrated up the faults from a greater depth, and not from the adjacent rock.

## Summary

The observations presented support the contention that the same fault can act as a conduit and barrier to fluid flow, and have anisotropic flow properties. The cyclic nature of the fluid flow and associated mineralization which apparently seals the fault until further displacement ruptures the gouge and permits another cycle of fluid migration suggests that the mechanical properties of the fault are far from "steady state". The development of a cataclastic fabric within the gouge during episodes of movement are consistent with observations of faults where fluids either have not been present or have not deposited solutes. With mineralization, however, the gouge is completely changed and apparently is totally "healed" overprinting the brittle fabric, and becoming mechanically continuous. Therefore, subsequent displacement has to rupture the gouge, so the frictional properties are significantly altered. On faults where processes such as this occur, our ability to predict the mechanical properties are severely limited at the present as existing laboratory studies on constitutive relations have been predicated upon the gouge being in mechanical steady-state.

## REFERENCES


Byerlee, J. D., V. Mjachkin, R. Summers, and O. Voevoda, 1978, Structures developed in fault gouge during stable sliding and stick-slip, *Tectonophysics,* 44,161-171.

Chester F. M. and J. M. Logan, 1987, Composite planar fabric of gouge from the Punchbowl fault, California, *J. Struct. Geol.*, 9, 621-634.

Decker, C. L., J. M. Logan, and V. B. Sisson, 1992, Mechanical behavior of faults in Somerset, U. K. due to fluid -rock interactions. *Proceedings of the Ussher Society.*

Feucht, L and J. M. Logan, 1990, The influence of chemically active fluids on the frictional properties of sandstone, *Tectonophysics,* 175,159-176.

Gatto, H. G., 1984, *The effect of various states of stress on the permeability of Berea Sandstone,* M. S. Thesis, Texas A&M Univ.

Higgs, N. G., 1981, *Mechanical properties of ultra-fine quartz, chlorite and bentonite in environments appropriate to upper-crustal earthquakes,* Ph. D. Diss., Texas A&M Univ.

Logan, J. M., M. Friedman, N. G. Higgs, C. Dengo and T. Shimamoto, 1979, Experimental studies of simulated gouge and their application to studies of natural fault zones, *Proc. of Conference VIII, Analysis of Fault Zones in Bedrock, U. S.* Geol. Survey, Menlo Park, Calif., 305-343.

Logan, J. M., N. G. Higgs, M. Friedman, 1981, Laboratory studies on natural gouge from the U.S.G.S. No. 1 well in the San Andreas Fault Zone, *Am Geophys. Union Monograph,* 24, 121-134.

Logan, J. M., 1992, The influence of fluid flow on the mechanical behavior of faults, *Proc. 33rd U. S. Symposium Rock Mechanics,* ed. J. R. Tillerson and W. R. Wawersik, 141-150.

Logan, J. M., C. A. Dengo, N. G. Higgs, Z. Z. Wang, 1992, Fabrics of experimental fault zones; their development and relationship to mechanical behavior, *in Fault Mechanics and Transport Properties of Rocks,* ed. Evans. B. and Wong. T-f, Academic Press, 33-68.

Rutter, E. H., R. H. Maddock, S. H. Hall, and S. H. White, 1986, Comparative microstructures of natural and experimentally produced clay bearing fault gouges, *Pure Appl. Geophys,* 124, 3-30.

Shimamoto, T., 1977, *Effects of fault gouge on the frictional properties of rocks: and experimental study,* Ph. D. Diss., Texas A&M Univ.

Wallace, R. E., and H. T. Wallace, 1986, Characteristics of faults and shear zones in deep mines, Pure and Appl Geop., v. 124, 107-126.

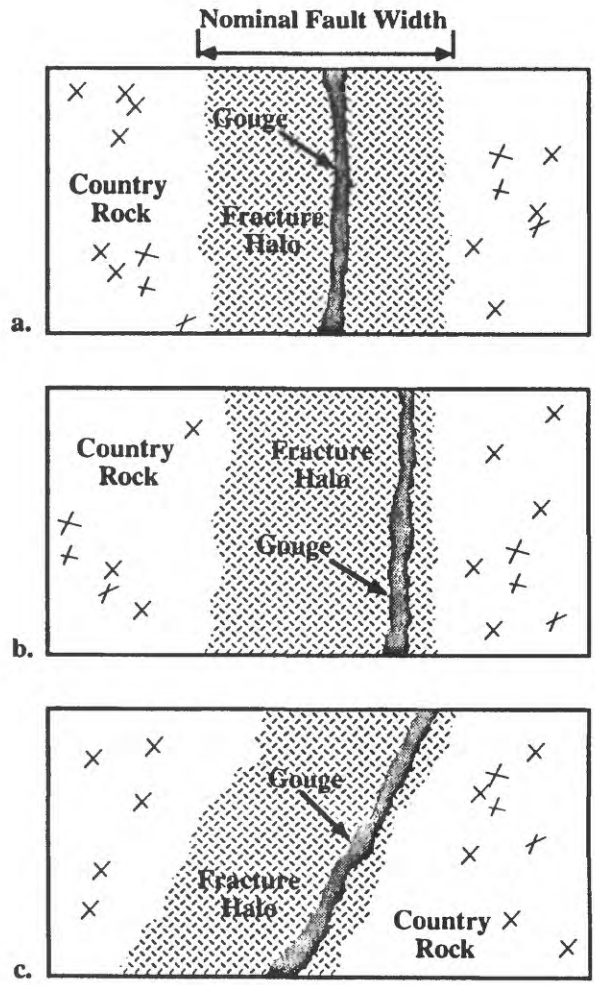



Figure 1. Diagrams showing characteristics common to many fault zones: country rock which may be fractured but less than a halo of highly fractured zone which in turn encases the gouge. The latter is highly sheared and comminuted material, which may be altered and/or mineralized. The gouge is often found at one boundary of the fractured halo as shown in (b) and (c). Note that the "nominal width" of the fault as mapped by the fracturing is often wider by ten times or more than the gouge where the displacement is localized.

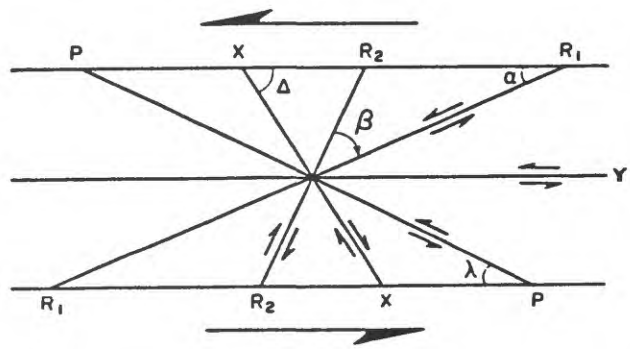



Figure 2. Fracture array and geometrical relationships defined by experiments on simulated fault gouge deformed in triaxial compression (From Logan et al., 1992).

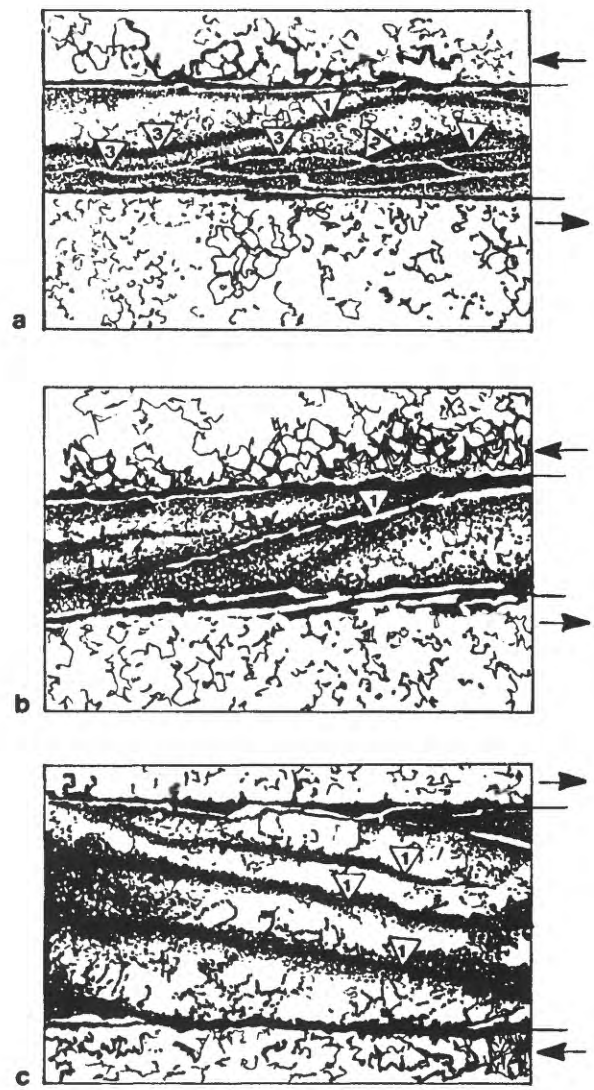



Figure 3. Sketches of simulated gouge zones deformed in the laboratory. Shear strains decrease from 12.7 in (a) to 6.7 in (b) to 3.1 in (c). Note the evolution of fabric with increasing shear strain, and eventual localization along the Y-fractures (numbered 3). R1 and P fractures are number 1 and 2 respectively. The more intense grain comminution indicated by the darker stippling along the fractures with the coarser-grained domains between them. It is inferred that these latter pods of higher permeability would be separated by lower permeability barriers along the fractures. ( From Logan et al., 1992).

Figure 4a. Fault zone illustrating localization of fluid flow as evidenced by calcite mineralization. Only the hanging wall is mineralized, with mineralization decreasing to zero within twenty meters of the fault. Normal relative displacement is only a few meters. Limestones and shales are offset by fault.

Figure 4b. View from cliff top at low tide. As the tide is about seven to eight meters the wave-cut terrace is exposed only part of the day. Faults seen can be traced along strike into cliff giving excellent 3-D exposure. As the faults are mineralized and mineralization is confined to a few meters of the fault, they are easily traced. Generally faults have normal relative displacement of only a few meters

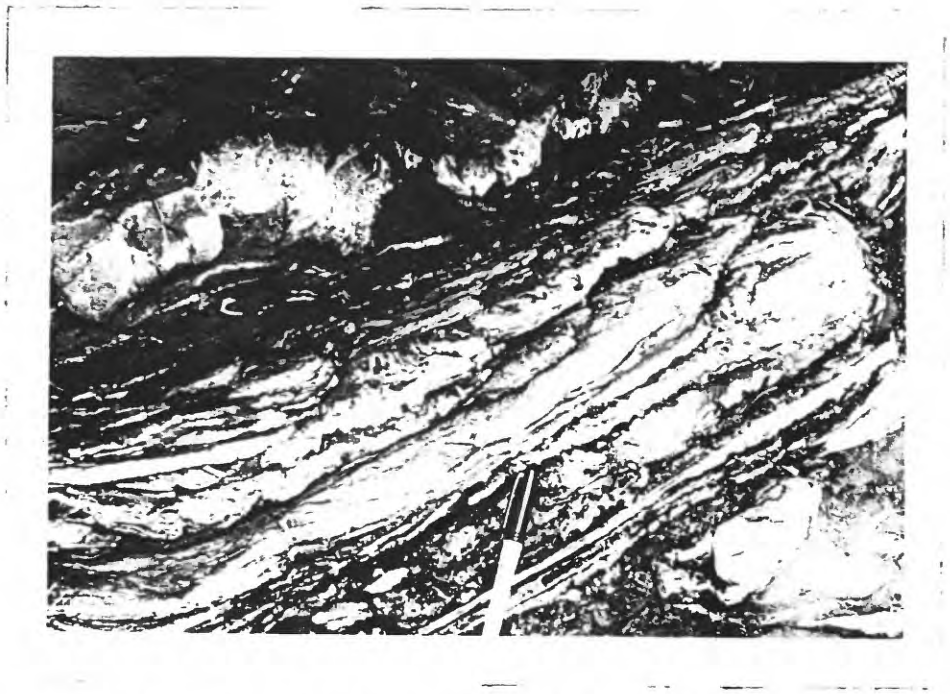


Figure 5. Fault gouge well mineralized with calcite and laminated with shale. Relative displacement is normal, with only tens of meters of displacement. Fabric elements in the R1 and Y orientations are visible. There is no evidence of mineralization a few centimeters away from the zone. Note brecciated limestone adjacent to upper boundary, which has been mineralized with calcite showing brittle character of the rock prior to mineralization.

Figure 6a. Normal fault with a few meters of relative displacement. Fault gouge is mineralized by calcite that is laminated with shale and while only a few centimeters thick it is continuous vertically for about 60 m and along strike for at least a kilometer. Lack of exposure limits tracing it further. Calcite-filled fractures in limestone of the hanging wall die out within a meter of the fault. Limestone and shale beds show reverse drag on both sides of the fault.

Figure 6b. Hanging wall of normal fault immediately to the gouge zone. Calcite has mineralized the fractures and the vug, although the latter remains unfilled, suggesting dilation of the fault zone during fluid flow.

Figure 7. Normal fault with calcite mineralization, but here only foot wall is mineralized. Relative displacement is only a few meters. Note, only selected fractures away from the fault zone are mineralized, and the permeability barrier to migration into the hanging wall must have been there before fluids migrated up the fault. Coin is about the diameter of a half dollar.

Figure 8. Normal fault shown in Figure 5 with mineralized gouge exposed along strike. The upper surface of the fault gouge is completely slickensided showing primarily dip-slip displacement. Note reverse drag in carbonates and shales of hanging wall. The fault zone as mineralized extends up dip for some twenty meters and along strike about thirty meters.

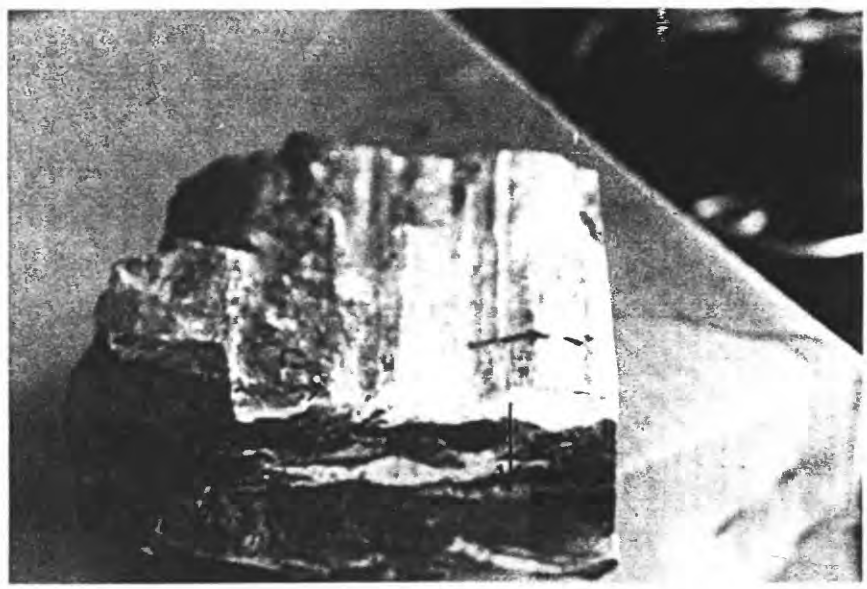


Figure 9. Hand sample taken from the fault shown in Figures 5 and 8. Note slickensides on upper surface which is the same surface shown in Figure 8. The laminated character of the calcite and clay-rich material in a Y-orientation is indicated by the arrow. It is in this interval that offset of markers in the outcrop indicate most of the displacement of the fault was accommodated. The missing upper block has moved toward the viewer. The sample is about 12 cm thick. Thin sections described in Figure 10 were taken from the right side of the sample.

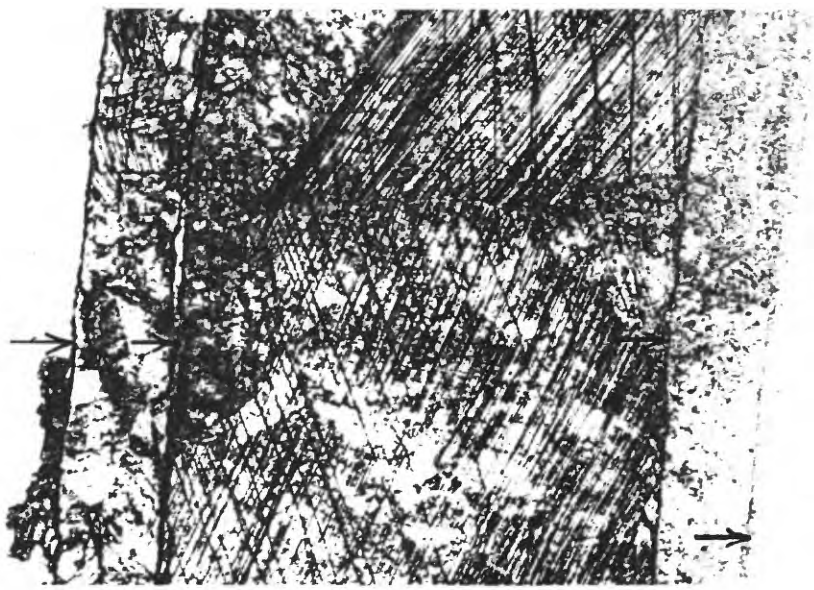


Figure 10. Photomicrographs of thin sections taken from sample shown in Figure 9. The left side of all photos is parallel to the upper and slickensided surface of Figures 5 and 8 with a left lateral sense of slip. The field of view in all is about 4 mm across. The series of micrographs progresses from (a) the slickensided surface at the edge of the fault zone towards the center of the fault (e). a. Photo at the slickensided surfaces of the fault which are marked by arrows. Note that there are three such surfaces. The coarse grained calcite displays a high density of twins. The high density suggests relatively high differential stress.

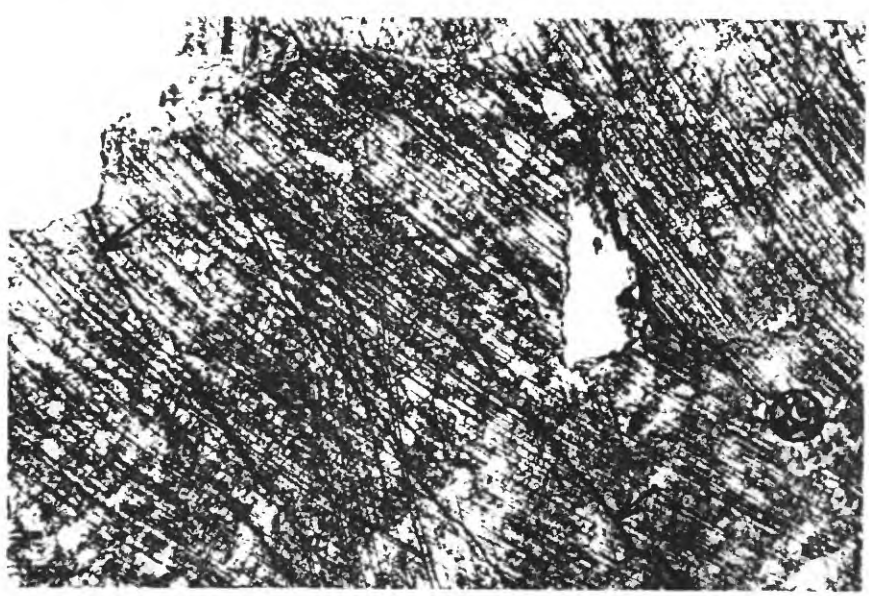


Figure 10b. Photo just to the right of (a), that is, closer to the center of the fault zone. The coarsely-crystalline calcite with twins have been strained as evidenced by the "bent" twins. These, in turn, have been subsequently fractured, with arrows indicating fractures.

Figure 10c. Photo just to the right of (b), now traversing towards middle of fault. The calcite is fractured with fragments separated by kerogen mixed with some clay minerals. On the right the vertical mineralized fractures, marked by arrows, are in the Y-orientation (Figure 2). The calcite is un-fractured, finer grained, and has very few twins relative to that on the left. This is the region of the well developed Y-surface found in the hand sample and in the outcrop where much of the displacement seems to have been localized

Figure 10d. Photo just to the right of (c). The Y-oriented mineralized fractures are evident. The cluster of inclined mineralized fractures on the right, indicated by arrows, are oriented in the P-orientation for normal slip. These and all other features are cut by a later set of fractures which have been mineralized; these are roughly horizontal in the photo and at high angles to the dip slip displacement. These have possibly been rotated and strained subsequent to their formation.

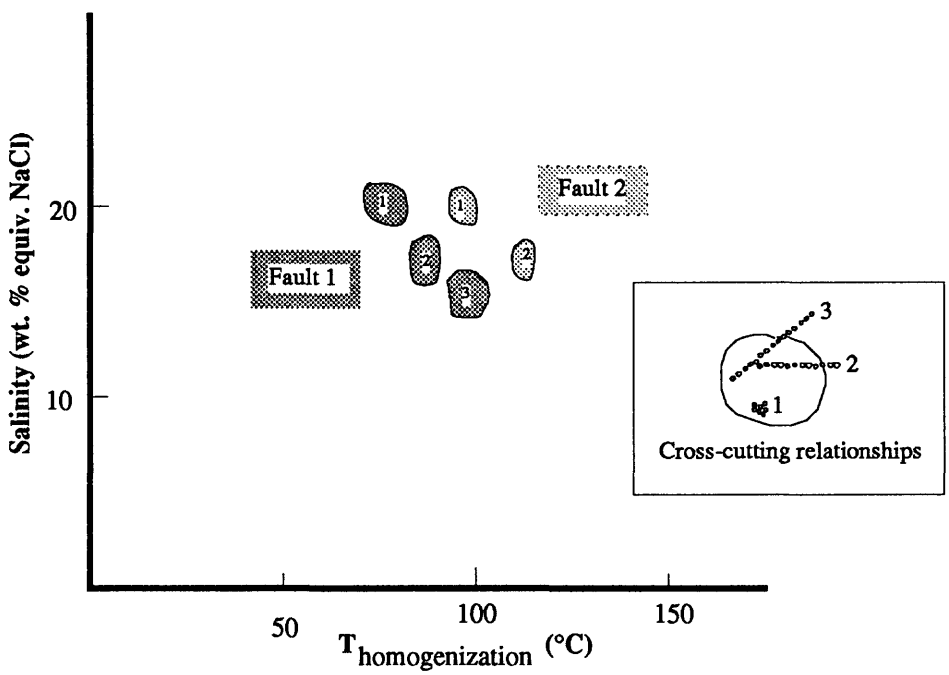



Figure 11. The cyclic mineralization and fluid flow demonstrated in textural observations of Figure 10 is supported by fluid inclusion samples taken from that fault and one other. The inset shows cross-cutting relationships among fluid inclusion clusters and trails used to determine three generations of fluids. The plot shows clusters of measurements, all from a single zone of mineralization within each of three faults, in terms of equivalent salinity (weight percent) and temperature of homogenization of trapped fluids. The size of the data clusters indicate the range of values measured from up to 200 measurements per sample and are color coded to the individual faults (modified from Decker, et al. 1993).

# Syn-deformational gold transport and deposition in brittle-ductile shear zones - some chaotic thoughts.

**R. W. Henley and M.A. Etheridge**

Etheridge and Henley - Geoscience Consultants, PO Box 3778, Manuka ACT 2603, Australia

## Abstract

Mineral deposits provide a record of crustal scale syn-deformational fluid flow events in the mid and upper crust. Yet while advances have been made in respect of understanding, for example, fluid pressure cycling in response to stress induced porosity and permeability changes at the outcrop scale, a number of assumptions are inherent in contemporary models of the larger scale fluid flow controls and particularly in estimation of the actual masses of fluid that are involved. For example, chemical and isotope tracers of fluid flow (such as gold) provide information only on the cumulative masses of fluid involved and not on the real mass of fluid which may be recycled perhaps through stress dilational processes. In this paper, dynamic fluid flow controls are addressed from the grain to crustal scales with emphasis on time variant permeability, the interaction of crack and porous media flow on mass transfer through shear zones, the coupling of chemical and deformational phenomena, and far-from-equilibrium chemical transport and deposition processes.

---

## Introduction

Modern crustal-scale shear/fault[1] zones generate large earthquakes. Their uplifted and eroded fossil equivalents host many of the world's major gold deposits. The principal connection between these two phenomena is, as pointed out by Sibson (1989), syn-deformational fluid flow within the shear zones from one crustal regime to another. Despite, however, a broad recognition of the role and behaviour of fluids in shear/fault zones, we know little of the details of how they relate to the processes of fault movement and earthquake generation, or the scales of the hydrothermal systems involved in their formation.

We do know, from consistent series of observations, that most of the major anomalous accumulations of metals in the crust, that are represented by ore deposits[2], are the consequence of episodic fluid flux along shear/fault zones in what were once tectonically active regions of the crust. The style, metal suite and magnitude of the deposits which are formed in these settings are the product of chemical processes that are coupled with fluid flow processes in and around shear/fault zones. The huge Pb-Zn SEDEX-style and Cu deposits of Mt Isa (Queensland) and major gold deposits such as Homestake, South

---

1 Much of the discussion in this paper focusses on mineral deposits and fluid behaviour in mid-crustal, largely ductile shear zones. However, it also relates to deposits in higher level, largely brittle fault zones, so that we use the term 'shear zone' here to include all crustal levels.

2 Mineral deposits, rather than rare highly randomised quirks of nature, as is so commonly assumed-even by exploration geologists, are the 'stochastic' , quite abundant, consequences of syn-deformational fluid flow. We often forget that the term "ore" reflects economic rather than geologic criteria, and that the two are only related through a series of probability terms (including the temporal variability of interest rates at the Bundesbank), perhaps analogous to the probability terms which are inherent in our understanding of sub-atomic processes.

Dakota, the Golden Mile (Kalgoorlie, Western Australia), Yellowknife, (Canada), and many others, are notable examples. Epithermal and VMS-style deposits, higher in the crust, are similarly associated with active and reactivated regional-scale brittle fracture systems in terrestrial and submarine settings respectively. We also know, from detailed studies, that many of these deposits have evolved through more than one period of episodic fluid flow.

The potential of any shear/fault array to develop economically viable mineral deposits, is a complex function of the magnitude of the cumulative fluid flux within it and the "reactivity" of the host rocks toward the fluid phase as it evolves within the shear zone. In addition to the coupled chemical processes involved, this is therefore a function of its permeability, the fluid pressure gradient[3], and the availability and supply of fluid.

The majority of current interpretive models of fluid flow in shear zones are variants of simple, single pass, steady state, pipe models (Figure 1a). Convection of fluid has also been proposed as a means of conserving and recycling smaller fluid volumes, but in their turn such models fail to address important questions of how fluid moves down through lithostatic gradients or how permeability is maintained in ductile crust.

These models represent highly simplified accounts of the reality of shear-zone mineralisation. Mineral deposits in shear zones are expressions of complexity. That is what we see. We do not see the consequences of reversible deformational, chemical or fluid flow processes. We therefore must carefully reconsider any model which accepts as its basis either steady state or equilibrium paradigms, or the stringencies imposed by assumptions of homogeneity.

For the objectives of the Red Book forum, in this paper we highlight some assumptions that are built into current models regarding the length and time scales of fluid flux and chemical phenomena in shear zones. The paper represents a stage in our thinking about the coupling of fluid flow and chemical processes in shear zones. In it, we raise a number of problems that we perceive in the present understanding of shear zone mineralisation, but intend to solve none.

## Shear zone hosted gold deposits - some boundary constraints

Shear zone hosted gold deposits, such as those along the Mother Lode(California), whether economic to mine or not, are the fingerprint of syn-deformational fluid flow in the earth's crust. The characteristics of this so-called "mesothermal" style of deposit have been summarised elsewhere (see, for example, Hodgson, 1989, and Kerrich, 1989 for recent reviews). The metals held within such mineral deposits act as tracers of the lithospheric processes involved both in the shear/fault zone, and the crustal regimes which act as sources of fluid, metals and associated components such as sulfur and arsenic.

Gold, which has a crustal abundance of only a few ppb (Crockett., 1991), has chemical properties which make it highly sensitive to changing temperature, fluid pressure and chemical conditions. These make it an ideal tracer of large scale fluid flow phenomena in the crust. In the individual deposits which form along shear zones in dilational and reactive sites, gold concentrations are highly variable (cf Koch and Link, 1971). Mean concentrations of around 100 ppb are common with concentration ranging over some four orders of magnitude with pseudo-log normal distributions. In some cases the higher

---

3 Note that fluid flow occurs in response to potential energy gradients, or head. Understanding of the coupling of fluid flow and chemical processes requires that this gradient be represented along the flow path by defineable pressures in order to describe the chemical processes which occur using equattions of state, etc.

concentrations are mineable but most commonly mining is confined to settings where a second population of gold concentrations, with a higher mean concentration, complements a lower grade population. Quartz-gold veins which crosscut the ductile fabric of shear zones are one common example of this. In some cases they may represent discrete deformational or chemical "happenings" within the shear zone during its evolution, but in others may represent quite separate evolutionary stages superimposed on existing shear zones.

Mineral assemblages record information about the P-T-X characteristics of the fluid in the ore depositional regime, the magnitude of the chemical fluxes involved, and the coupled chemical and physical response of the fluid-rock couple throughout the life time of the shear zone. This information provides the basis of viable attempts to model the complexity of shear zone evolution, provided that we read the record correctly and do not impose artificial boundary constraints on the interpretation. Mineral assemblages also record information about the source regime(s) for water, gold, $CO_2$, sulfur, radiogenic isotopes, potassium, etc, which therefore act as tracers of fluid flow processes. Kerrich (1989) and others have concluded, from radiogenic and stable isotope and chemical tracer data, that multiple crustal and mantle lithosphere sources contribute to the geochemistry of altered shear zones. Structural data inside shear zones also suggest that the chemical fluxes indicated by the tracers occurred within rocks during the development of both ductile fabrics and cross-cutting brittle fractures. None of these observations provides information on the scale of phenomena involved.

**Fluid and metal sources**

Since the solubility of gold in hydrothermal fluids is well constrained by experiment (e.g. Seward, 1991) and by observation at a maximum of about 10ppb, it is evident that a 1 million ounce (31 tonnes) gold deposit represents a fluid flux in excess of $3km^3$. Based on the average gold concentration of crustal rocks (1-2 ppb), this indicates a minimum source regime volume for the gold alone, of some $30km^3$ . These are minimum estimates because economic cut-offs, determine that only about 30% of the gold accumulated in shear zones is usually accounted for in determining the mineable resource. A million ounce ore deposit therefore more likely the record of the passage of more than $10km^3$ of fluid. The unknowns in estimating total transported gold and therefore source rock volumes for the fluid and metal are essentially those relating to process efficiency.

Individual shear zones commonly host multiple gold deposits a number of which may contain mineable gold in excess of 1 million ounces. Figure 2 provides an example of an Archean shear zone system in Western Australia which hosts a number of deposits whose known cumulative mineable gold content is about 21 tonnes (Watkins and Hickman, 1990). This is small in comparison with giant shear zone hosted deposits such as the Golden Mile at Kalgoorlie with its total resource in excess of 1200 tonnes. Presumably the host rock/structural setting of the Golden Mile provided an ideal, highly efficient sink for gold during fluid migration in the controlling shear system. Taking this larger resource, the minimum fluid volume required for gold transport in a major shear zone is well in excess of $100km^3$. Assuming a nominal 2 wt% of fluid released by metamorphic dehydration reactions in the source regime, or directly from magmatic processes, the required source rock volume for the fluid, on a one time single pass basis is about $2000km^3$. Since the shear zone array shown as the exemplar has a minimum length of 1000km, the fluid source rock volume within the upper 5km of the exposed crust corresponds to a mean width for the shear zone of less than 0.5km. Although we know little of the mean and range of widths of unmineralised segments in such shear zones, open pit gold operations have, over the last few years, shown that the mineralised segments of shear zones are generally much wider

than 0.5km, and that many comprise an anastomosing array of shears over a width of several kilometres. We infer therefore that the source rock volumes required to produce given sets of gold deposits, may be regarded as of similar magnitude to the volumes of the shear zones that host them. The above crude estimates suggest that neither the rock volumes for derivation of either the fluid or the gold are significant problems to resolve with respect to understanding of mineralisation and fluid flow in shear zone arrays.

We wish to here highlight a key assumption that is buried within the volume calculations. They assume that the fluid passes through the shear zone only once. We can see no justification for this ad hoc assumption since no chemical or isotopic tracer at our disposal can discriminate between single or cumulative incremental processes during mineralisation in shear zones. The chemical and isotopic tracer signatures of large scale fluid flow on shear zones are cumulative effects which may arise as readily through the passage of large fluid volumes as through multiple recycling of smaller packets of fluid through the same volume of deforming rock, provided that the fluid packet can interact with a larger source volume between cycles.

## Fluid Flow and Chemical Processes in Mineralised Shear Zones

### Fluid Flow-Length scales

On the macro- or crustal scale, shear zones act as guides for fluid flow in response to fluid pressure gradients (i.e."head"). Since permeability and transmissivity are related via the cross section of the fluid flux, the long range, long term fluid flow in shear zones may therefore be considered to be dominated by flow through the low permeability rock matrix of the shear zone. However, in addition to being within zones of ductile shear, the singular common characteristic of all shear zone-hosted gold deposits is that brittle failure, and crack-seal veining are developed repeatedly during the ambient ductile deformation process. The resultant veins in some cases constitute high grade ore zones. In order to understand fluid flow in shear zones, it is therefore required that we address the question as to whether the transient, high transmissivity cracks or the grain scale permeability of the ductile shear zone itself control the time-integrated fluid flux through the shear zone.

Short term, or transient, fluid pressure cycling and fluid flow behaviour during crack-seal/shearing events, have been discussed in detail by Cox et al. (1990) and by Bouillier and Robert (1992), and correlated with seismogenic phenomena in the modern, active crust (e.g. Sibson, 1989 ). We can infer from these studies that the short range behaviour of shear zones is dominated by transient fluid pressure changes. At the micro-, or grain scale, Fischer and Paterson (1989) have shown that during ductile deformation, stress-induced dilatancy occurs by microfracturing, and this therefore transiently increases the connected porosity and permeability of the stressed rock by an order of magnitude or more at high ambient fluid pressures. This stress dilatancy effect has been considered elsewhere (e.g. Cox et al., 1990) but only recently demonstrated experimentally at high temperatures (Fischer and Paterson, op. cit., Cox, this volume).

The contrast between the long range, long term steady state and short range transient concepts emphasises the need to consider what length scale is needed for realistic phenomenological modelling of the coupling of fluid flow and chemical processes in shear zones.

The characteristic length scale for the long range permeable medium behaviour of shear zones may be taken as $10^0$ to $> 10^2$ km (Figure 2). That for stress dilatancy is unknown but, based on the scale of phenomena seen as discrete mineralised bodies of rock within

individual deposits in shear zones, is likely to be of the order of $10^2$ m. These observations give us a wide choice of length scales to base modelling of flow in shear zones.

The permeability of a fracture or fracture array is several orders of magnitude higher than that of the dilated matrix. In turn, the latter is at least an order of magnitude greater than the host rock outside the shear zone. Realistic process models for fluid flow and chemical dispersion in shear zones therefore must be formulated in terms of the permeability of both the crack and the matrix, rather than one or other, or some convenient effective permeability parameter. Such process models fall into a class of "factorizable fruit cake" mathematical models (eg McNabb, 1978). Roughly translated, these models address the question as to whether the properties of a fruit cake are determined by the fruit or by the cake mix, or, of course, by both (Figure 3). The modelling problem, in the case of a shear zone, is formulated by considering that, at any instant, the shear zone is divided into blocks of characteristic length, $L_i$ , such that over distances greater than $L_i$, the system behaves as an isotropic porous medium whose permeability is greater than that of the rock matrix. Thus the fracture network provides a wide range of path lengths for fluid flow such that the composition of fluid at any point in the system is dependent on (a) the relative proportion of long and short path lengths, and (b) the relative rates at which fluid bleeds into or from the matrix into the fracture system.

It is the fluid pressure changes which occur over this short range length scale which drive the mixing of wallrock-reacted and unreacted, or modified, source fluid and therefore initiate the geochemical processes responsible for rock alteration and ore formation.

## Chemical processes

Most popular models of the chemical processes responsible for gold mineralisation take as their starting assumption that the fluid phase is always at, or very close to saturation with respect to the pH, redox and ligand concentrations of the fluid. There is no basis for such a limiting assumption. Chemical equilibrium or reversibility is commonly assumed in building process models for mineral deposits despite the fact that they imply incremental deposition of gold along the flow path and the impossibility of forming mineral deposits.

The observations that gold occurs at concentrations several orders of magnitude above its crustal abundance, and that in mineral deposits it is associated with discrete structural sites or reactive rock units, suggest that natural processes such as metal transport and deposition, require "far-from equilibrium," irreversible chemical process paradigms as the basis for chemical modelling (see for example Prigogine and Stengers, 1985, Berger and Henley, in prep.). Mineral deposits occur when the separate kinetics of all of the fluid and fluid-rock reactions which are triggered by deformational events, become complementary through positive feedback mechanisms. Mineral deposits are the remnants of "dissipative" geochemical and fluid dynamic structures, in a sense "strange attractors" in both the chaotic and societal senses.

A common example is the association of major gold deposits with active shears or faults crosscutting carbonaceous or ferrous-iron rich lithologies. In this environment carbon dioxide rich fluid penetrates the shear zone host rock, where fluid-rock reactions convert much of the carbon dioxide to methane. Back flow of reacted fluid into the shear zone, perhaps due to stress release by fracturing, produces mixtures of wallrock-reacted and source fluids, which in turn release hydrogen from the now excess methane. This on-site evolution of hydrogen, lowers the solubility of gold, regardless of its complexation, such that gold is deposited. Positive feedback processes may occur in the system to catalyze further deposition. The recognition of this type of chemical drive mechanism for gold

deposition allows that the gold concentration of the original shear zone fluid may have been well below saturation during transport from its source regime.

## Fluid sources revisited

Analysis of mineralised shear zones indicates that transient effects such as stress dilatancy and crack-seal vein formation, may control the movement of fluid in shear zones and therefore control the mixing of components from distinct source regimes above, below and lateral to active shear zones (Fig. 1b). A shear zone will draw its fluid from as many and whatever sources are available to it at any instant and which lie within the fluctuating fluid pressure field along the shear zone. This model of fluid flow in shear zone arrays also allows us to reconsider the estimates made of fluid volume requirements for mineralisation, which, with the single pass assumption, indicated that well over 2000km$^3$ of fluid was required.

By contrast, shear-zones, and their contained crack-seal veins, might be regarded as natural oscillators rather than single-pass pump/valve systems[4] provided that we removed the single pass assumption. For arguments sake, we propose that only a fraction of the above volume is required to produce shear-zone hosted gold deposits, but that it is recycled or shunted around through and around the shear zone in response to changing fluid pressure gradients due to imposed stress, transient stress dilatancy effects and the heterogeneity of the system. What new paradigms might arise from this proposition?

Hammond and Nisbet (1990) have shown that a number of deformational stages can be recognised in each of the Archean greenstone belts Western Australia, with shear-zone mineralisation occurring at each stage along with granite magmatism. They also showed that mineralisation occurred both in early extensional shear zones and later strike-slip style shear zones. Some of these late shear zones may be reactivated components of earlier extensional or thrust arrays. Together they comprise an evolving pattern of inter-related brittle-ductile shear arrays.

If we viewed therefore the structural and tectonic history of such Archean volcanic belts on the scale shown in Figure 2, we would expect to see a heterogeneous, time dependent pattern of temperatures and fluid pressures. If we then proposed that the gold and fluid sources were coupled, either through metamorphic processes or through release of fluid from crystallising magma, we would see a discontinuous pattern of fluid evolution and pulses of fluid migration at discrete heat flow or deformation sites along the connected shear zone array. Stress dilatancy effects might also draw fluid along the shear zone toward these sites, leading to a complex space-time pattern of deformation, fluid production, fluid migration and gold transport and deposition. Subsequent magmatism or deformation may trigger new patterns of fluid migration and evolve new centres of mineralisation and

---

4 The behaviour of discharging geothermal wells is instructive although not strictly analogous to shear zone fluid flow behaviour. Some wells discharge fluids whose measured enthalpy reflects that of vapour pressure fluid at the reservoir temperature. Others discharge fluid mixtures of anomalous enthalpy due to mixing of fluids from discrete reservoirs. Excess enthalpy wells may be shown through mass and energy balance models, supported by chemical and isotopic data, that they develop by either preferential admixture of steam (with $CO_2$) from depressurised aquifers, or by enthalpy increase due to conductive heat gain from wallrocks. The longevity of the latter as production wells is limited but they may recover after shut down. At Kawerau (New Zealand) for example, the excess enthalpy effect was such as to increase lead concentrations from a few ppm to several thousand and allow surface precipitation of lead oxychlorides (B. Christenson, pers. comm.1985). Other wells behave as oscillators with two or more aquifers supplying fluids in proportions reflecting their own fluid pressure differentials and that of the flowing well.

reactivate older mineralised sites in the three-dimensional network Figuratively it would be like the sequencing of lights in an evening citiscape.

An intriguing analogue of such a crustal scale natural process is to be found in modern understanding of brain activity. Thus neurons stimulated (thermal or stress nodes) in one section of the brain "fire" and release chemical messengers to adjacent or distal neurons (depositional nodes), and these in turn transmit chemical signals further through the neural network. We do not want to push the analogy too far, but find it interesting to contemplate whether memory effects might also operate in shear zone systems!

### Remobilisation or syndeformational redistribution of elements

Whilst the concept of "remobilisation" of ore components has been controversial, it becomes almost axiomatic in this conserved fluid concept that metals will become redistributed in response to changing fluid pressure gradients within an evolving shear zone because of (a) fractured porous media effects, and (b) changing fluid saturation with time.

Consider a shear zone as it is initiated in the lower or middle crust (Figure 4a). At this point, the gold distribution, along and within it, is at the average crustal abundance (~1 ppb). During the operation of the shear zone, as discussed above, fluid flow is controlled by the grain scale permeability of the shear zone (itself a function of strain rate and progressive alteration history), and transiently by the higher local permeabilities induced by stress. The accumulated effect of fluid flow through reactive lithologies in the shear zone leads progressively to the development of a low grade population of gold along and within the zone (Figure 4b). Now consider the transient brittle failure events which lead to the development of crack-seal vein arrays. In each transient event, fluid pressure increase due to strain drives fluid into the low permeability shear zone matrix and host rock, which now contains gold at an average of 100 times its crustal abundance and mostly residing outside silicate and sulfide crystal lattices. Interaction of fluid with this matrix evolves the fluid chemistry toward some local equilibrium state. Failure and stress release to form high permeability fractures then develops a new transient fluid pressure gradient regime where the pressure in the matrix exceeds that in the cracks so that back flow occurs of the evolved fluid to mix with introduced unmodified fluid. Since the matrix fluid has evolved in an environment containing essentially free gold, its gold concentration is now close to the solubility of gold at the ambient matrix fluid conditions of sulfur, hydrogen, proton, etc activities. Mixing of matrix and shear fluids results in a fluid which is far from equilibrium with respect to its end-members leading, as described above, to gold deposition through a chain of reactions dominated by those producing hydrogen[5]. Inevitably this leads to syn-deformational redistribution of gold from a steady state syn-deformational "protore" environment to develop higher, localised gold contents in the crack-seal veins(Figure 4c).

## Concluding Statement

The intent of this paper has been to highlight a number of unresolved problems which relate to the flow of fluid in active shear/fault systems. The most pervasive of the

---

5 Note that the mineral parageneses developed in this manner are a function of local stress cycling rather than of changing bulk fluid composition as is commonly assumed in interpreting mineral parageneses. Moreover, this model does not demand that common parageneses encountered along shear zones were developed simultaneously, only that certain fluid rock interactivities were repeated. Mineral parageneses are the expression of the strange attractor response of fluid chemistry to transient, deformation-triggered fluid mixing-or unmixing.

paradigms which influence our thinking about shear zones, are those relating to the magnitude of the fluid volume, and the restraints imposed by equilibrium chemical modelling. These constraints on our understanding of fluid flow, and mineral deposition, in active shear zones may be released by considering the interplay of length and time scales in shear zone evolution.

It is possible that popular pump-valve models for fluid flow and mineral deposition in and around shear zones are too simplistic. Although they have usefully focussed geological attention on transient fluid flow phenomena in shear zones, they should be replaced by more complex models including the dynamics of fractured porous media, far-from equilibrium chemical transport and depositional scenarios, and the recycling of more limited fluid volumes than are dictated by single pass models.

## Acknowledgment

We would like to thank the organisers of the Red Book meeting for their invitation which triggered this chaotic paper and attempted to control its kinetics. We also thank Drs. S. Cox[2] for their valuable discussions and comments, and B. R. Berger for helping to crystallise some thought processes.

## References

Bouiller, A.-M., and Robert, F., 1992, J. Structural Geology, v.14, p. 161-180.

Cox, S.F., Etheridge, M.A., and Wall, V.J., 1990, in Robert, F., Sheahan, P.A. and Green, S.B.(eds), Greenstone Gold and Crustal Evolution, Geological Assoc. Canada, p. 46-53.

Crockett, J.H., 1991, Distribution of gold in the Earth's crust, in Foster, R.P.(ed.), Gold Metallogeny and Exploration, Blackie, Glasgow, p. 1-36.

Fischer, G.J., and Paterson, M.S., 1989, J. Geophysical research, v. 94, p. 17607-17617.

Hodgson, C.J., Patterns of Mineralization, in J.T. Bursnall (ed.), Mineralization and Shear Zones, Geol. Assoc. Canada, p. 51-89

Kerrich, R.W., 1989, in J.T. Bursnall (ed.), Mineralization and Shear Zones, Geol. Assoc. Canada, p. 129 - 198.

Koch, G.S. and Link, R.F., 1971, Statistical Analysis of Geological Data, Wiley, N.Y., p. 363-389.

McNabb, A., 1978, N.Z. Maths Soc, First Australasian Mathematics Convention, Invited Address.

McNabb, A., and Henley, R.W., Proc. Geothermal Workshop, Geothermal Institute, University of Auckland, p. 68-71.

Prigogine, Ilya and Stengers, Isabelle, 1985, Order out of chaos, Flamingo-Harper, Glasgow, 348p.

Seward, T.M., 1991, The hydrothermal chemistry of gold, in Foster, R.P.(ed.), Gold Metallogeny and Exploration, Blackie, Glasgow, p. 37-62.

Sibson, R.H., 1989, J. Structural Geology, v.11, P. 1-14.

Watkins, K.P. and Hickman, A.H., 1990, Geological evolution and mineralization of the Murchison province, Western Australia, Geological Survey of Western Australia, Bulletin 137, 267p.

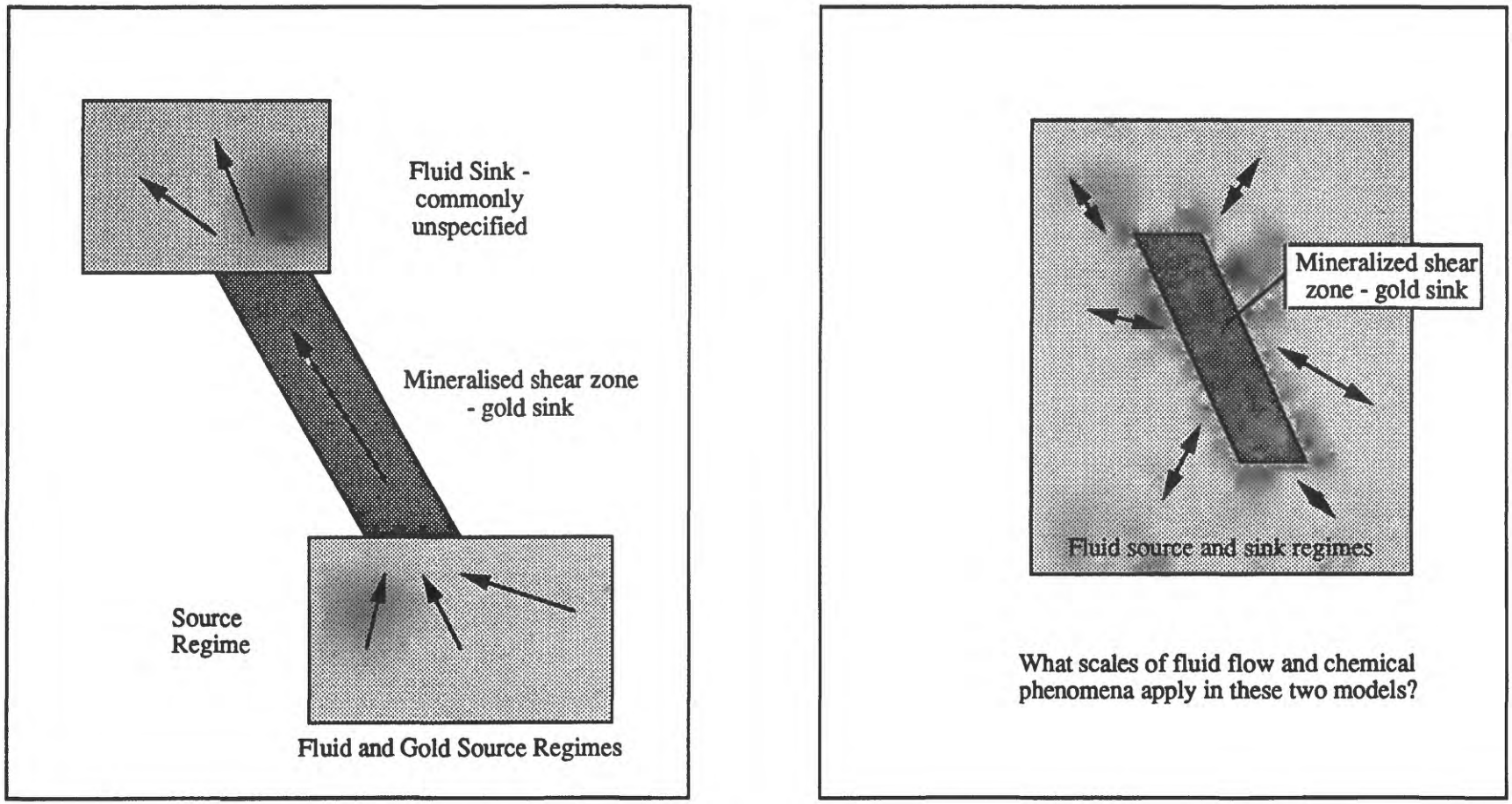



Figure 1. Simplistic cartoons illustrating (a) a typical single pass flow model ("fault-valve model") for a shear-zone hosted gold deposit, and (b) an alternate model("oscillator") model in which source and sink regimes are identical. The latter incorporates process features of the fault-valve

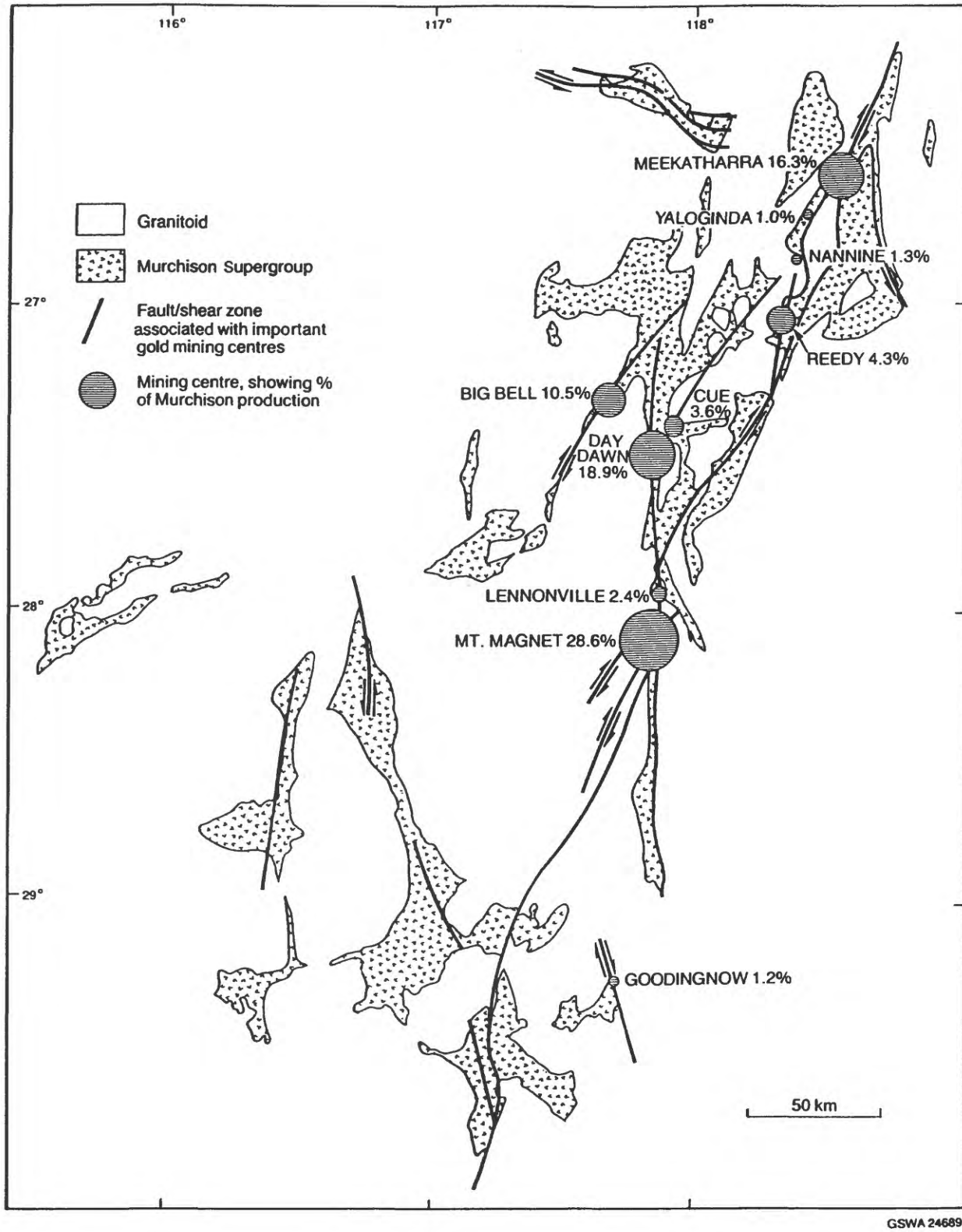



Figure 2. Distribution of the major shear zones and major gold deposits in the Archean Murchison greenstone belt in Western Australia. Only the major gold mining centres are shown together with their gold production to 1988 in terms of a percentage of the 228 tonnes total gold production from the belt. . Minor prospects and mines are not shown but are more widely distributed through the shear zone array (Reproduced with permission from Watkins and Hickman, 1990).

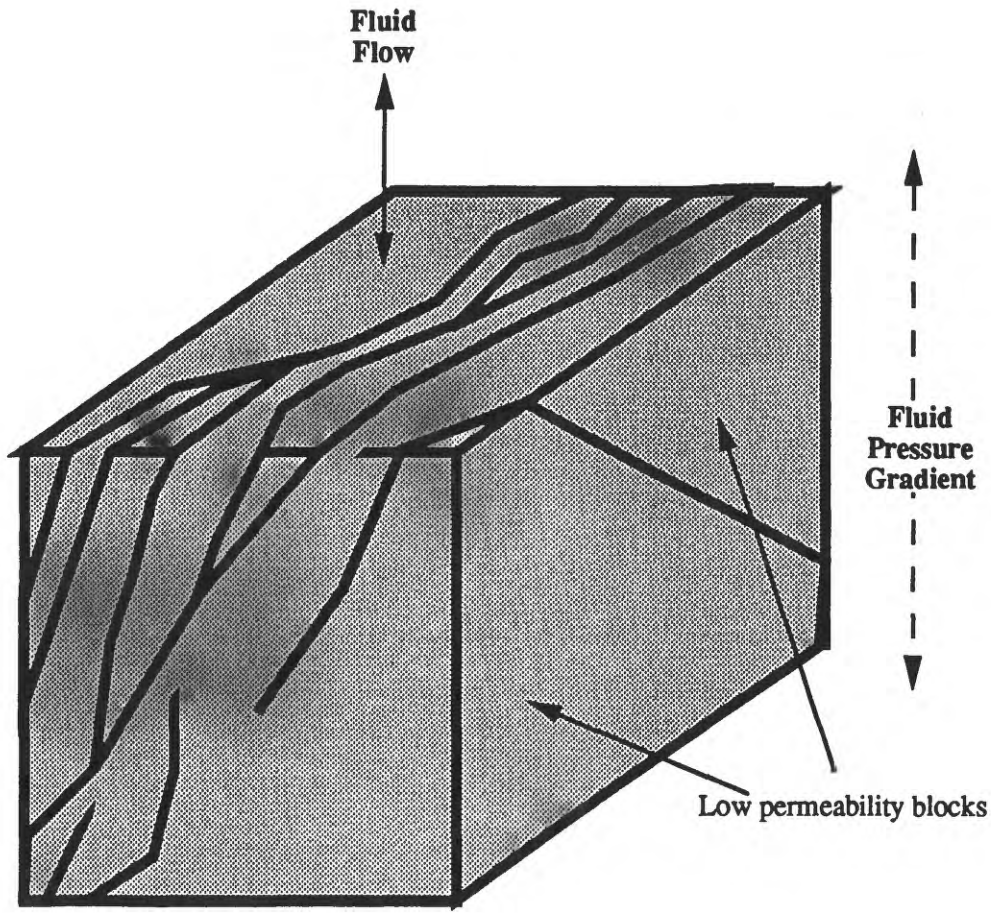



Figure 3. In a fractured porous me dium of scale, l, is the bulk flow of fluid under any steady state or transient fl uid pressure gradient regime, determined by the porous blocks, the cracks or both?

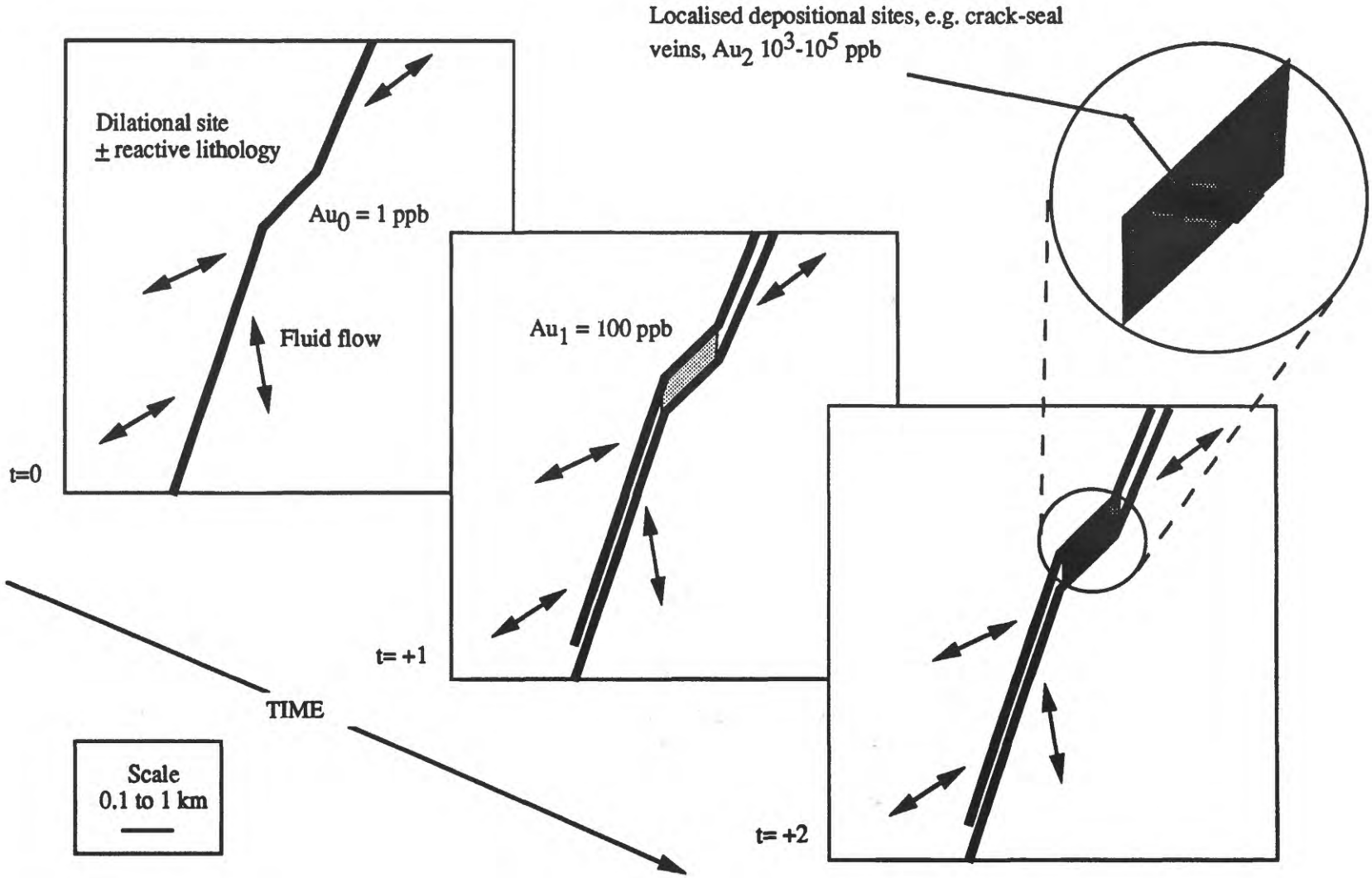



Figure 4. A schematic representation of the evolution of a shear zone hosted gold deposit emphasizing the progressive upgrading of the local abundance of gold as a consequence of coupled chemical and brittle-ductile processes(see text). The oscillator mechanism flip-flops between states t=1 and t=2 allowing syn-deformational development of the low and high concentration gold populations and their spatial distributions. The schema applies to a wide range of shear zone settings and is not confined to the simple jog dilatancy site used here for illustration.

# PERMEABILITY STRUCTURE OF A THRUST FAULT

Craig B. Forster
Dept. of Geology & Geophysics, University of Utah, Salt Lake City, UT 84108

James V. Goddard and James P. Evans
Dept. of Geology, Utah State University, Logan, UT 84322

## Abstract

A clear picture of the permeability structure of faults is required in order to provide input to mathematical models of fluid flow processes within and near faults. To this end we have characterized the architecture of a thrust fault through detailed mapping of fabrics, fractures, and alteration of faults found in a crystalline thrust sheet. The resulting information is consistent with a conceptual model for the fault that comprises three primary components; intact protolith, a fractured damaged zone, and a gouge core. Core samples obtained from blocks of exhumed fault rock were selected to include representative samples of each component of the fault zone. Permeability tests performed on 31 cores reveal that, at low confining pressures, the permeability of the damage zone ($10^{-1}$ to 10 md) is approximately two orders of magnitude greater than that of the adjacent protolith and gouge ($10^{-5}$ to 1 md). Overall the permeability of the fault samples range over 6 orders of magnitude. Testing a subset of the samples as a function of confining pressure suggests that, despite differences in absolute value, the permeability of the gouge and damaged zone samples responds in a similar fashion to changes in confining pressure. At the scale of the samples tested our results indicate that distinct hydromechanical properties may be associated with each of the three primary components of a fault zone.

## Introduction

Developing and revising conceptual models that describe fault evolution through space and time requires an improved understanding of the permeability structure of faults. For example, Rice (1992) makes specific assumptions regarding the variation of fault rock permeability as a function of confining pressure that, in turn, impact his calculation of pore pressures within the San Andreas Fault. Data obtained from laboratory tests performed on samples of fault rock cored in deep drilling projects and from *in situ* tests in faults at depths of 1 to 5 km, however, are rare (Smith et al., 1991). To make best use of this sparse dataset, we recommend shallow *in situ* testing and laboratory testing of exhumed fault rocks. When coupled with detailed mapping of fractures, fabric, and rock alteration the results of shallow test results should provide important clues regarding the permeability structure of faults. As a first step in pursuing this approach we report permeability data obtained from laboratory tests of exhumed fault rock obtained from the Washakie fault system of northwestern Wyoming. A more detailed discussion of the background for this study is provided by Goddard (1992).

## Geologic Setting

Rocks tested in this study are obtained from the East Fork thrusts in the Washakie Ranges of northwestern Wyoming. The East Fork fault system is in the hanging wall of a cystalline thrust sheet. The exhumed fault zones formed at a depth of 4 to 7 km, have granite and gneiss in both footwall and hanging wall, and experienced 2 to 4 km of slip (Evans, in press). Quartz veins, iron-oxide-coated fractures, and syntectonically altered feldspars indicate that fluids penetrated the fault during slip (Evans, 1990).

The East Fork fault system comprises a set of relatively simple structures with well-defined histories and abundant evidence for syntectonic fluid flow. Aside from regional geothermal gradients and intra-fault friction, no other heat sources were likely present during fault evolution. Thus, syntectonic fluid flow was likely driven by the elevated topography that developed during thrusting (Forster and Evans, 1991). Possible sources for the fluids includes meteoric water and formation waters trapped in the crystalline thrust sheet and the underlying sedimentary rocks.

Faults found within the East Fork fault system are either narrow, well-defined faults or wide zones of fault gouge. Narrow faults comprising well-foliated, indurated silica-cemented, fine-grained gouge and cataclasite are spaced 1 cm to 1 m apart. Wide (up to 6 m wide), large displacment (0.1 to 4 km of slip) faults were responsible for emplacement and imbrication of the thrust sheet. These faults consist of zones of fine-grained gouge and cataclasite, with crude to well-defined foliation, that separate lozenges of fractured and altered protolith.

The fault zones are internally zoned. These zones can be roughly classified as: 1) unaltered protolith, 2) a fractured and faulted damaged zone, and 3) indurated gouge and cataclasite zones. A conceptual model for the architecture of a typical fault zone is shown in Figure 1. Here the protolith is stripped away leaving the pinching and swelling pods of gouge and damaged zone. A scale is not included in Figure 1 because the geometry incorporated in this conceptual model applies across all scales; from micro- to megascale.

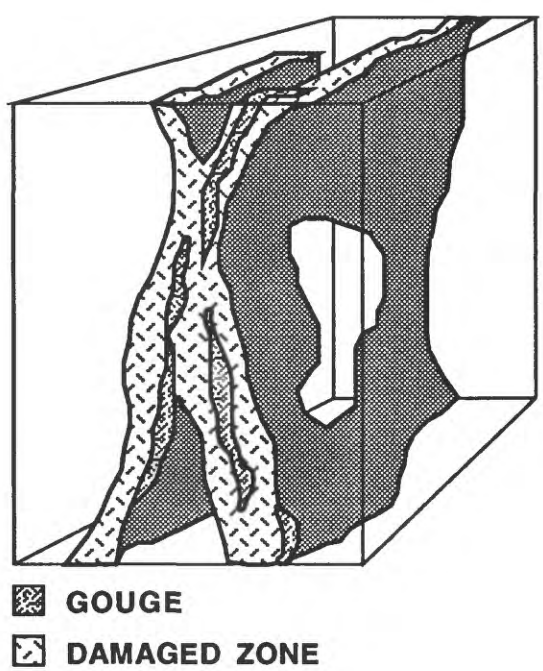



Figure 1. Conceptual model of a fault zone with the protolith stripped away to reveal the external geometry of the gouge and damaged zones.

In the East Fork fault system the unaltered granitic protolith consists of coarse-grained microcline, plagioclase, quartz, and biotite with accessory Fe-Ti oxides. The damaged zone composition is identical to the protolith, except that small amounts of kaolinite and iron oxide are found in the fractures. Subsidiary faults in the damaged zone lie at angles of 1 to 30 degrees to the fault plane. Within the gouge zone microcline, plagioclase, quartz, and biotite are altered to sericitized feldspar, kaolinite, and iron oxide staining.

The density of fractures and subsidiary faults increases toward the center of the fault zone. Fracture patterns perpendicular to the slip direction and the plane of the fault are roughly orthogonal. Fracture patterns parallel to the slip direction and perpendicular to the plane of the fault, however, are approximately diamond shaped. Quartz veins are pervasive at both macroscopic and microscopic scales in the damaged zone. Layers of iron oxides and magnetite found in the quartz veins suggest several episodes of fluid flow.

Brittle deformation and comminution of feldspars were the dominant deformation mechanisms (Evans, 1988 and 1990). Feldspars deformed by comminution from extension and shear fractures primarily along cleavage planes. Quartz deformed primarily by brittle fracture. Syntectonic alteration of feldspars to clay minerals resulted in the formation of foliated cataclasites. Foliation in the fault gouge and cataclasite is defined by alignment of phyllosilicates, segregation and elongation of minerals, fractures and faults which are nearly parallel to the fault boundaries, and seams of secondary and opaque minerals.

Increased slip was accompanied by alteration of feldspars to phyllosilicates and reduction of grain size (Evans, 1988 and 1990). Wide, indurated gouge zones commonly display an irregular, wavy texture at the hand-sample scale. Fault boundaries range from sharp to gradational. Grooved, polished and slickenlined surfaces are common.

## Permeability Testing

Figure 2 illustrates how 31 core samples were selected, at a variety of orientations with respect to the slip plane, to provide test results representing the primary components (protolith, gouge, damaged zone) of a typical fault zone. Two core sizes were tested; 2.54 cm (1 inch) and 5.08 cm (2 inch) diameter. Sample lengths are generally at least twice the sample diameter. The smaller cores (19 in number) were tested with water as the permeant at a single effective confining pressure of 3.5 MPa using a steady state method. The larger cores (12 in number) were tested with a gas permeant using a transient pressure pulse method. Each of the larger samples was tested at a series of confining pressures ranging from 2 to 47 MPa. All tests were performed by Terra Tek, Inc. in their Salt Lake City laboratory.

The results of permeability tests performed at $P_{eff}$ of about 4 MPa are shown in Figure 3 as a function of location within the fault zone. Core diameters are indicated by the boxes (5.08 cm) and circles (2.54 cm). The larger diameter, lower permeability protolith samples are unweathered and unfractured rocks. The higher permeability of the smaller protolith cores is attributed to the effects of weathering. Although both gouge and damaged zone rocks often contain partially healed and throughgoing fractures, the gouge cores are distinguished from damaged zone cores by the presence of significant stringers of cataclasite within the gouge rocks.

Two features of the results shown in Figure 3 are of interest. First, within each category (protolith, damaged zone, gouge) a wide range of k was obtained; on the order of 2 to 3 orders of magnitude. Second, the damaged zone appears to have a consistently higher k than that of the adjacent gouge and protolith. Overall, fault rock permeabilities range from less than $10^{-5}$ md to greater than 10 md; a range of more than 6 orders of magnitude.

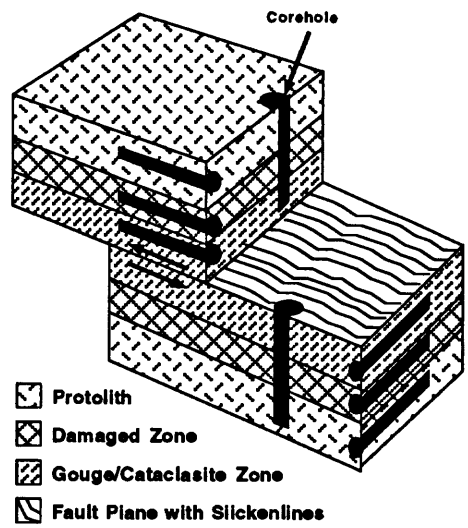



Figure 2. Block diagram of a typical fault zone showing the generalized orientations and locations of cores extracted from exhumed fault rocks. Cores are 2.54 and 5.08 cm in diameter.

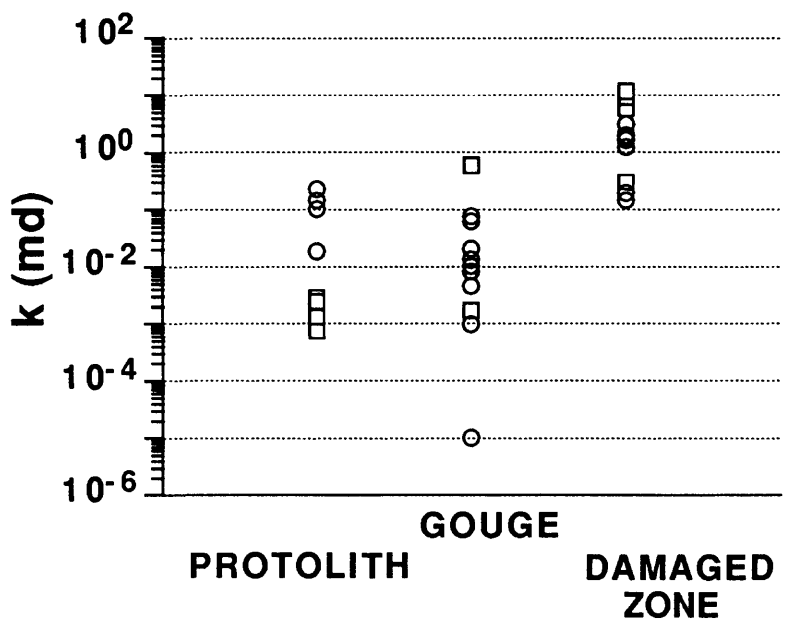



Figure 3. Laboratory-derived permeabilities obtained for each fault rock category at effective confining pressures of about 4 MPa. Core diameters are indicated by boxes (5.08 cm) and circles (2.54 cm).

The wide range of k found in our laboratory test results is consistent with the range in k values obtained from detailed *in situ* tests performed in shallow thrust faults found in both Canada (Davison and Kozak, 1988) and Sweden (Andersson et al.,1991) and with the results of other laboratory tests ($10^{-6}$ to $10^{3}$ md) [see Smith et al., 1990]. The *in situ* tests, however, yield consistently higher k values (ranging from $10^{-2}$ to $10^{5}$ md). The lower permeability of the core samples likely reflects the fact that relatively large, *in situ* fractures are not represented in the core samples.

A normalized permeability ($k_{norm}$) is defined to aid in comparing permeability results obtained as a function of confining pressure ($P_{eff}$) for each of the fault rock categories:

$$k_{norm} = \frac{k_i}{k_1}$$

where $k_1$ is the value of k obtained at $P_{eff}$ = 4 MPa and $k_i$ is the value obtained at the $i^{th}$ level of confining pressure. $P_{eff}$ is computed as applied confining pressure minus applied pore pressure. Values of $k_{norm}$ are plotted as a function of $P_{eff}$ in Figure 4.

Figure 4 illustrates the variation in normalized permeability with effective confining pressure obtained for each category of fault rock. Data obtained for the 4 samples of protolith (Figure 4a) are fitted with an exponential model of form:

$$k_{norm} = m_0 \exp [P_{eff}\, m_1]$$

where $m_0$ and $m_1$ are fitting parameters that vary from 1.426 to 1.509 for $m_0$ and from -0.170 t o -0.230 for $m_1$. Values of k were not obtained at effective confining pressures greater than 10 MPa because the lower limit of the permeability test method was reached at this value of $P_{eff}$.

Data obtained for 1 gouge sample (6-91-3) and 4 damaged zone samples are shown in Figure 4b and fitted with a logarithmic model of form:

$$k_{norm} = m_0 + m_1 \log [P_{eff}]$$

The fitting parameters vary from 1.165 to 1.283 for $m_0$ and from -0.570 to -0.813 for $m_1$. In almost all cases, the values of R2 (parameter describing "goodness of fit") fall within the range 0.98 to 0.99 for the curves illustrated in Figure 4.

The higher permeability of the gouge and damaged zone samples (Figure 4) allowed testing to continue to confining pressures in excess of 30 MPa. Comparing Figures 4a and 4b suggests that the protolith undergoes a greater change in k with changes in confining pressure than that of the gouge and damaged zones. This, in turn, implies that a greater "stiffness" is imparted to the gouge and damaged zones by the presence of throughgoing fractures that are absent from the protolith samples.

The models obtained in this analysis are similar to other empirical and theoretical models derived to describe variations in the permeability of both whole and fractured rock as a function of confining pressure (e.g., Brace et al., 1968; Gangi, 1978; Tsang and Witherspoon, 1981). We are currently exploring the theoretical basis for the observed results with a view to providing a more physically-based model for extrapolating our results to greater depths within faults. It is of particular interest to establish differences in the hydromechanical behaviour of distinct elements of fault zones (e.g., gouge, protolith, and damage zones).

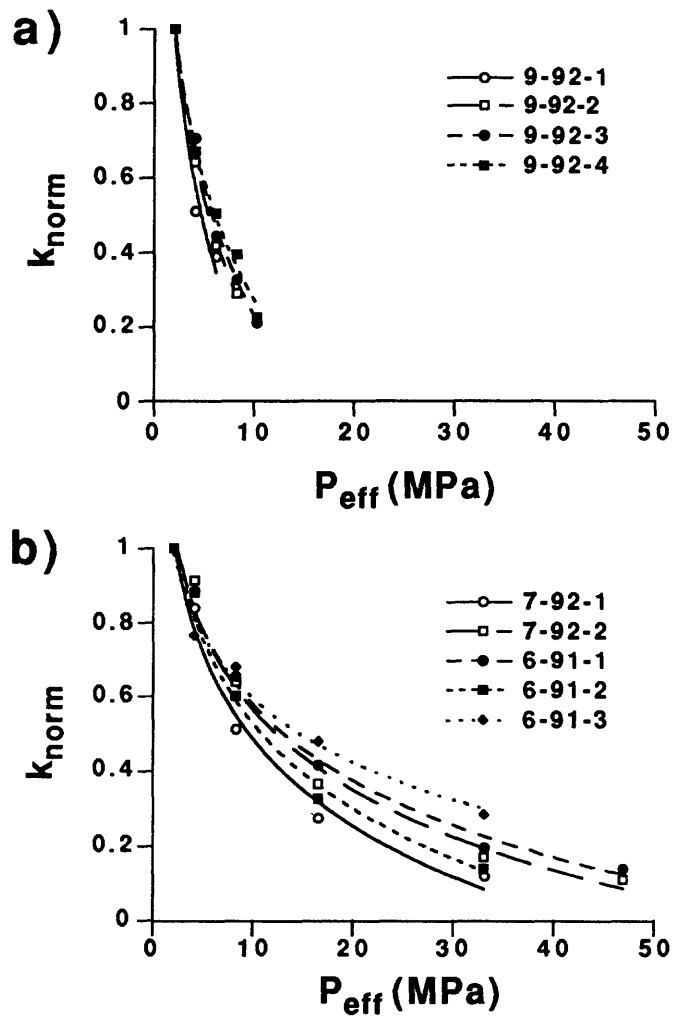



Figure 4. Normalized permeability ($k_{norm}$) as a function of effective confining pressure ($P_{eff}$); a) protolith data with exponential fit, and b) gouge and damaged zone data with logarithmic fit. Filled symbols represent gouge samples in (b).

Using the fitted relationships found for the protolith, gouge, and damaged zone rocks (Figures 4a and 4b) we develop generalized models for the variation of k with $P_{eff}$ that might be representative of each element of the thrust fault. This is accomplished by averaging the fitting parameters obtained from the curve fits shown in Figure 4a (protolith) and Figure 4b (gouge and damaged zone) to define characteristic relationships for each element of the thrust fault. Although the same logarithmic model can be fitted to both gouge and damaged zone rocks, each component of the fault has a distinctly different range of permeability (Figure 4). Consequently, the values of k computed at $P_{eff} = 0$ differ for each of the gouge, protolith, and damage zones (Figure 5). The maximum and mininum curves shown in Figure 5 are computed using the maximum and minimum values of k obtained in testing the 5.08 cm cores.

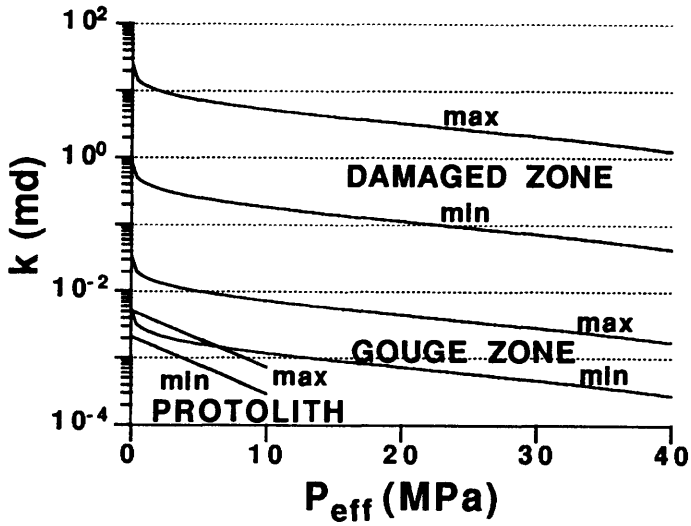



Figure 5. Extrapolated permeability (k) as a function of confining pressure ($P_{eff}$) for damaged zone, gouge zone, and protolith. The range of variation of k within each rock category is given by showing the generalized logarithmic (damaged and gouge zones) and exponential (protolith) functions obtained using the maximum and minimum values of k measured in the laboratory testing program.


## Summary

The generalized models developed for each element of the fault zone provide insight into the way that k declines with increasing $P_{eff}$. Computing the difference between the k inferred at a specified confining pressure and the value of k found on the same curve for Peff = 0 may indicate the possible magnitude of permeability change anticipated when pore pressures rise, during a seismic event, from hydrostatic (high $P_{eff}$) to lithostatic (low $P_{eff}$). For example, in a transition from hydrostatic to lithostatic pore pressure at $P_{eff}$ = 10 MPa, k for the protolith may increase by about 1 order of magnitude while k for the gouge and damaged zones may increase only by a factor of 5. The apparent variation in hydromechanical properties of the gouge, protolith, and damaged zones suggests that careful mapping of permeability-stress relationships within each element of a fault will aid in constructing plausible hydromechanical models for faults and aid in extending the sparse data obtained from deep *in situ* tests in fault zones.



*Acknowledgements.* Funding was provided by National Science Foundation grant #EAR-90050897, Utah State University research grants, and the University of Utah.

## References


Andersson, J.-E., Ekman, L., Nordqvist, R., and Winberg, A., 1991. Hydraulic testing and modelling of a low-angle fracture zone at Finssjön, Sweden. Jour. of Hydrol., v. 126. pp. 45-77.

Brace, W.F., Walsh, J.B., and W.T. Frangos, 1968. Permeability of granite under high pressure. J. Geophys. Res, v. 6., pp. 2225-2236.

Davison, , C.C., and C.Y. Wang, 1988. Hydrogeologic characteristics of major fracture zones in a large granite batholith of the Canadian shield. Proc. 4th Canadian-American Conf. on Hydrogeology, Banff, Canada, June, 1988.

Evans, J.P., in press. Deformation mechanisms and kinematics of a crystalline-cored thrust sheet: The Washakie thrust system, Wyoming. in "Basement Deformation in Rocky Mountain Foreland Structures" (Schmidt, C.J., Chase, R., and Erslev, E.A., eds.), Geol. Amer. Spec. Paper 280.

Evans, J.P., 1990. Textures and deformation mechanisms and the role of fluids in cataclastically deformed granitic rocks, in "Deformation Mechanisms, Rheology, and Tectonics" (Knipe, R.J., and E. Rutter, eds.), Geol. Soc. London Spec. Publ n0. 54, pp. 29-39.

Evans, J.P., 1988. Deformation mechanisms in granitic rocks at shallow crustal levels. Jour. Struct. Geol. v. 10, pp 437-443.

Forster, C.B. and J.P. Evans, 1991. Hydrogeology of fault zones and crystalline thrust sheets: Results of field and modeling studies, Geophys. Res. Lett., v. 18, 979-982.

Gangi, A.F., 1978. Variation of whole and fractured porous rock permeability with confining pressure. Int. J. Rock Mech. and Min. Sci., v. 15, pp. 249-257.

Goddard, J.V., 1993. Internal Deformation, Evolution and Fluid Flow in Basement-Involved Thrust Faults, Northwestern Wyoming. M.S. Thesis, Dept. of Geology, Utah State University, 134 pp.

Rice, J.R., 1992. Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault. in "Fault Mechanics and Transport Properties in Rocks" (Evans and Wong, eds.), Academic Press, pp. 475-503.

Smith, L., Forster, C.B., and J.P. Evans, 1990. Interaction of fault zones, fluid flow and heat transfer at the basin scale. in "Hydrogeology of Low Permeability Environments" (S.P. Neuman and I. Neretnieks, eds.), Int. Assoc. Hydrol. Sci., Selected Papers in Hydrogeology, v. 2, pp. 41-67, Verlag.

Tsang, Y.W. and P.A. Witherspoon, 1981. Hydromechanical behavior of a deformable rock fracture subject to normal stress. Jour. of Geophys. Res., v. 86, pp. 9287-9298.

# Hill fault/fracture meshes as migration conduits for overpressured fluids

Richard H. Sibson

*Department of Geology, University of Otago*
*P.O. Box 56, Dunedin, New Zealand*

**Abstract:** Hill (1977) interpreted earthquake swarm activity as resulting from the passage of fluids through an interlinked 'honeycomb' mesh comprising minor faults and extension fractures. There is field evidence for the development of such fault/fracture meshes over a broad range of scales, and for their role as channels for fluid migration. They form effective fluid conduits when the level of fluid pressure approaches the least compressive stress, $\sigma_3$ (i.e. $\sigma_3' = (\sigma_3 - P_f) \rightarrow 0$). Swarm activity at other than very shallow depths is thus probably diagnostic of localized fluid overpressures. Though such meshes may develop in all tectonic regimes, their formation is especially favored by interlayered competent and incompetent strata in extensional settings where they develop as extensional "chimneys" or "fuzzy" normal faults. The passage of large fluid volumes through such meshes, presumably from the intermittent breaching of overpressured fluid compartments, may lead to the formation of breccias with maximum clast size related to the thickness of competent layers. The hypothesis that earthquake swarms at other than very shallow depths are diagnostic of fluid overpressures is testable by drilling.

## Fluid pressure, faulting and extension fracturing

The character of fracturing at depth in the earth's crust is determined by rock material properties, the level of differential stress, $(\sigma_1 - \sigma_3)$, and the fluid pressure, $P_f$. The fluid pressure level can be expressed in terms of the pore-fluid factor, $\lambda_v = P_f/\sigma_v$, where $\sigma_v$ is the overburden pressure. Hydrostatic fluid pressures ($\lambda_v \sim 0.4$) prevail where interconnected pore space and fracture systems are freely linked through to the earth's surface. However, at depths greater than a few kilometres in deforming crust, fluid pressures commonly rise well above hydrostatic values and may approach lithostatic values ($\lambda_v \rightarrow 1$) (Fyfe *et al.* 1978). In a rock mass under triaxial stress (principal compressive stresses, $\sigma_1 > \sigma_2 > \sigma_3$), extension fractures form perpendicular to $\sigma_3$ by hydraulic fracturing when:

$$P_f = \sigma_3 + T \qquad (1)$$

provided $(\sigma_1 - \sigma_3) < 4T$, where T is the long-term tensile strength of the rock (Secor 1965). For sedimentary rocks, T is commonly of the order 1-10 MPa (Etheridge, 1983). When $(\sigma_1 - \sigma_3) > 4T$, shear failure occurs in accordance with the Coulomb criterion:

$$\tau = C + \mu_i \sigma_n' = C + \mu_i(\sigma_n - P_f) \qquad (2)$$

where C, the cohesive rock strength, and $\mu_i$, the coefficient of internal friction, are material properties, and $\tau$ and $\sigma_n$ are, respectively, the shear and normal stresses on the failure plane. In

homogeneous, isotropic, intact rock, faults therefore initate as Coulomb shears containing the $\sigma_2$ axis and lying at $\pm$ 20-30° to $\sigma_1$ (Anderson 1951).

## Stress controls on permeability

Stress-controlled features affecting rock permeability include brittle faults, microcracks, extension fractures, and stylolite seams. Their relationship to a stress field with principal compressive stresses $\sigma_1 > \sigma_2 > \sigma_3$ in homogeneous, isotropic rock is as shown in Figure 1, which illustrates the range of fault/fracture/stylolite combinations that may develop in different circumstances. Macroscopic extension fractures formed by natural hydraulic fracturing and microcracks developing by grain impingement have a preferential orientation perpendicular to $\sigma_3$. Impermeable stylolites develop as anticracks in planes perpendicular to $\sigma_1$ (Fletcher & Pollard 1981). The different components may combine in various fault/fracture/stylolite meshes as shown, but the full range of components need not be present in all cases. Faults and fractures have the potential to impart directional permeability to the rock mass, but may also become choked with hydrothermal deposits and/or clay-rich gouge and alteration products (e.g. Hooper, 1991).

Existing anisotropic permeability in layered rock sequences may therefore be either enhanced or counteracted, depending on the type of stress field prevailing and the dominant attitude of the layering (Fig. 2). In the case of predominantly horizontal layering, and for the more common situation where fracture-enhanced permeability predominates over impermable stylolitic bands, bedding-parallel permeability is enhanced in a compressional tectonic regime, while trans-bedding permeability is enhanced in extensional and strike-slip regimes. Note that permeability is also enhanced out of the plane of the 2-D section illustrated, parallel to the intersections of the faults and fractures. This is especially important in strike-slip settings, where greatly enhanced vertical permeability may result from the localised development of mesh structure in dilational jogs and bends along strike-slip faults (Sibson, 1986).

## Fluid expulsion through Hill fault/fracture meshes

Hill (1977) showed that high b-value earthquake swarms could develop from the passage of hydrothermal or magmatic fluids through a 'honeycomb mesh' of interlinked shear and extensional fractures, triggering a cascade of minor shear ruptures (Figs. 1 & 3). The requisite condition for the hydrofracturing component, $P_f > \sigma_3$, may be achieved under hydrostatic fluid pressures close to the earth's surface in areas of extensional and strike-slip tectonics (Fig. 4), but hydraulic fractures may also develop at depth in the crust when fluid pressure approaches the lithostatic load ($\lambda_v \rightarrow 1$) (Sibson, 1981). Fault-fracture meshes are therefore most likely to form in extensional and strike-slip settings as fluid pressures at least equal to the lithostatic load (i.e. $\lambda_v \geq 1$) are required at all depths for hydrofracturing in compressional stress regimes when $\sigma_v = \sigma_3$. Development in extensional and strike-slip settings is also favoured by the fact that individual vertical hydro-fractures can only extend over limited depth intervals (Secor & Pollard 1975). Fluid migration through such meshes would induce $\Delta P_f$ fluctuations of the order of the rock's tensile strength (1-10 MPa).

Hill fault/fracture meshes are particularly likely to form in well-layered rock sequences undergoing horizontal extension with $\sigma_3$ horizontal and $\sigma_1$ vertical (Fig. 3), because the brittle failure mode may be dictated by the tensile strength of alternating rock layers. For example, in a

typical sequence of alternating sandstones and shales with $T_{sst} > T_{sh}$, the shales may fail by development of brittle Coulomb shears ($(\sigma_1 - \sigma_3) > 4T$), perhaps in conjugate sets, while the sandstone fails by hydraulic fracturing ($(\sigma_1 - \sigma_3) < 4T$), to give rise to subvertical extensional "chimneys" or "fuzzy" normal faults as illustrated in Fig. 3. As a consequence of heterogeneity and anisotropy within the rock mass, natural meshes are likely to take on a more irregular 3-D form, especially in transtensional environments where $\sigma_2 \approx \sigma_3$. The ability of such fault/fracture meshes to transmit fluids is highly sensitive to the value of the effective least compressive stress, $\sigma_3' = (\sigma_3 - P_f)$, but becomes large when $\sigma_3' \rightarrow 0$. The passage of large fluid volumes may lead to brecciation with rock fragments initially bounded by combinations of shear and extensional fractures, and maximum clast size governed by the thickness of the competent layers. Continued flow causes rotation and progressive comminution of the fragments in a fluidisation process, forming extensive tabular bodies of high-dilation breccia such as those described from the Monterey Shale formation of California (Redwine, 1981).

Fault/fracture-meshes developed in well-layered sedimentary sequences tend to form either as subvertical extensional "chimneys", or as "fuzzy" normal faults with an overall component of shear across the mesh (Fig. 3). In the Monterey Shales they have acted as major fluid expulsion structures, transporting large volumes of aqueous and hydrocarbon fluids across the bedding anisotropy, the relict mesh structure being preserved through cementation by silica, carbonate, and bitumen (Fig. 5).

## Breaching of overpressured compartments

Recent oil-field studies have demonstrated the widespread occurrence of fluid pressure compartments in sedimentary basins, and analysis of oil migration paths suggests that overpressured compartments become breached from time to time with episodic fluid discharge (Hunt 1990, Powley 1991). Though perhaps most widespread in compressional tectonic regimes, overpressuring is also known to occur in areas undergoing extensional faulting such as the Gulf Coast (Hunt, 1990; Hooper, 1991). Ruptures transecting suprahydrostatic gradients in fluid pressure and breaching impermeable seals give rise to *fault-valve* action, with upwards fluid discharge along the transient permeability of the fault zone and localized reversion towards a hydrostatic fluid pressure gradient.

Extensional breccia chimneys and fuzzy normal faults are therefore likely to have developed by the rupturing of seals to overpressured fluid compartments in extensional tectonic settings, leading to fluid expulsion and earthquake swarm activity. Given the extensive development of extensional mesh structures in the Monterey Shales exposed along the coast of the Santa Barbara Channel, it is interesting to note the present-day swarm activity within the channel (Sylvester *et al.*, 1970) and the existence of locally high fluid overpressures within adjacent sedimentary basins (e.g. $\lambda_v \sim 0.9$ at only 3 km depth within the Ventura Basin - Yeats, 1983). However, it should be noted that current deformation within the channel is predominantly compressional in character.

## Discussion

Meshes of interlinked faults and extension fractures of the kind envisaged in Hill's (1977) model for earthquake swarms can be recognised in outcrop. The presence of hydrothermal veining (and, sometimes, hydrocarbons) localised within the mesh structures testifies to their importance as

conduits for concentrated fluid migration. Taking account of the conditions for hydraulic extension fracturing, it appears that such mesh structures are only likely to develop under hydrostatic fluid pressures close to the Earth's surface in extensional/transtensional settings. At depths of more than a kilometre or so, swarm activity is probably dependent on fluid overpressures. In compressional settings, swarm activity at any depth likely denotes near-lithostatic levels of fluid pressure.

Earthquake swarms at depths of a few kilometres present definable targets for scientific drilling and other forms of geophysical investigation to test the hypothesis that swarm activity is associated with the migration of overpressured fluids. Penetration of swarms by drilling would allow direct measurement of fluid pressure levels, and sustained measurements might allow correlation of fluid pressure fluctuations with seismic activity.

## References


Anderson, E.M. 1951. *The Dynamics of Faulting* (2nd edn.), Oliver & Boyd, Edinburgh.

Etheridge, M.A. 1983. Differential stress magnitudes during regional deformation and metamorphism: upper bound imposed by tensile fracturing. *Geology* 11, 231-234.

Fletcher, R.C. & Pollard, D.D. 1981. Anticrack model for pressure solution surfaces. *Geology* 9, 419-424.

Fyfe, W.S., Price, N.J. & Thompson, A.B. 1978. *Fluids in the Earth's Crust.* Elsevier, Amsterdam, 383 pp.

Hill, D.P. 1977. A model for earthquake swarms. *J. Geophys. Res.* 82, 1347-1352.

Hooper, E.C.D. 1991. Fluid migration along growth faults in compacting sediments. *J. Petrol. Geol.* 14, 161-180.

Hunt, J.M. 1990. Generation and migration of petroleum from abnormally pressured fluid compartments. *Bull. Am. Assoc. Petrol. Geol.* 74, 1-12.

Powley, D.E. 1990. Pressures and hydrogeology in petroleum basins. *Earth Sci. Rev.* 29, 215-226.

Redwine, L. 1981. Hypothesis combining dilation, natural hydraulic fracturing, and dolomitization to explain petroleum reservoirs in Monterey Shale, Santa Maria area, California. *In*, Garrison, R.E., Douglas, R.G., Pisciatto, K.E., Isaacs, C.M. & Ingle, J.C. (eds.) The Monterey Formation and Related Siliceous Rocks of California, *Spec. Publ. Soc. Econ. Pal. Mineral.* 221-248.

Secor, D.T. 1965. Role of fluid pressure in jointing. *Am J. Sci.* 263, 633-646.

Secor, D.T. & Pollard, D.D. 1975. On the stability of open hydraulic fractures in the earth's crust. *Geophys. Res. Lett.* 2, 510-513.

Sibson R.H. 1981. Fluid flow accompanying faulting: field evidence and models. *Am. Geophys. Union Maurice Ewing Series* 4, 593-603.

Sibson, R.H. 1986. Earthquakes and lineament infrastructure. *Phil Trans. R. Soc. Lond.* A317, 63-79.

Sylvester, A.G., Smith, S.W. & Scholz, C.H. 1970. Earthquake swarm in the Santa Barbara Channel, California, 1968. *Bull. Seism. Soc. Am.* 58, 1047-1060.

Yeats, R.S. 1983. Large-scale Quaternary detachments in Ventura Basin, southern California. *J. Geophys. Res.* 88, 569-583.

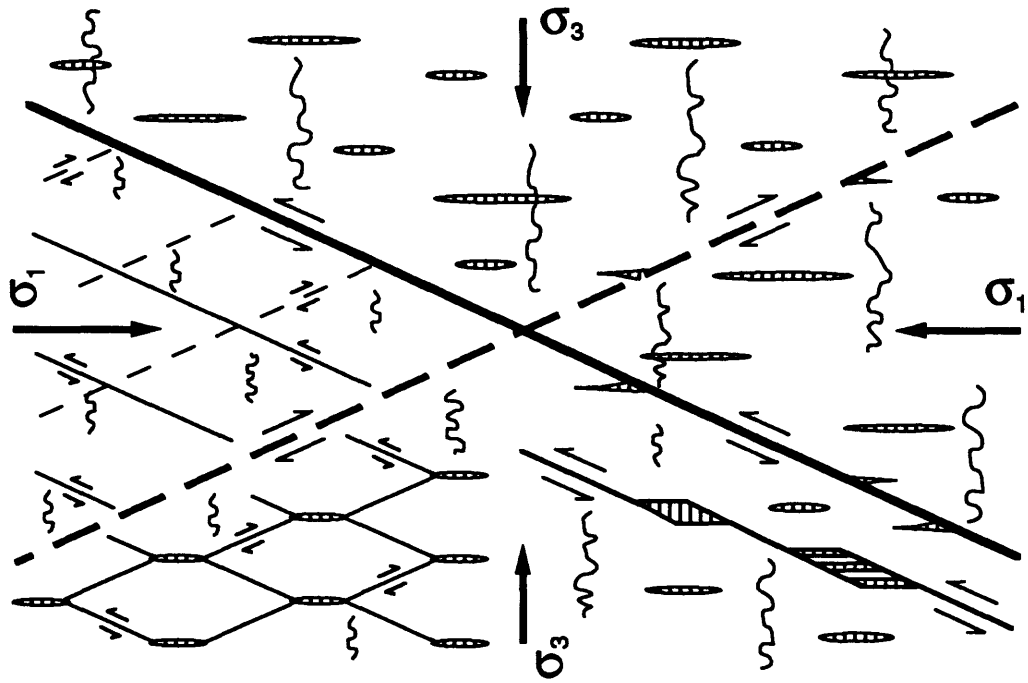



Figure 1 - Faults (Coulomb shears), extension fractures (hachured), and stylolites (squiggly lines) forming in a triaxial stress field, illustrating possible combinations of stress-controlled features which may affect permeability of the rock mass. A Hill fault/fracture mesh is shown in the lower left. Diagram represents a thrust fault stress regime when upright, a normal fault regime when viewed sideways, and a strike-slip regime in plan view.

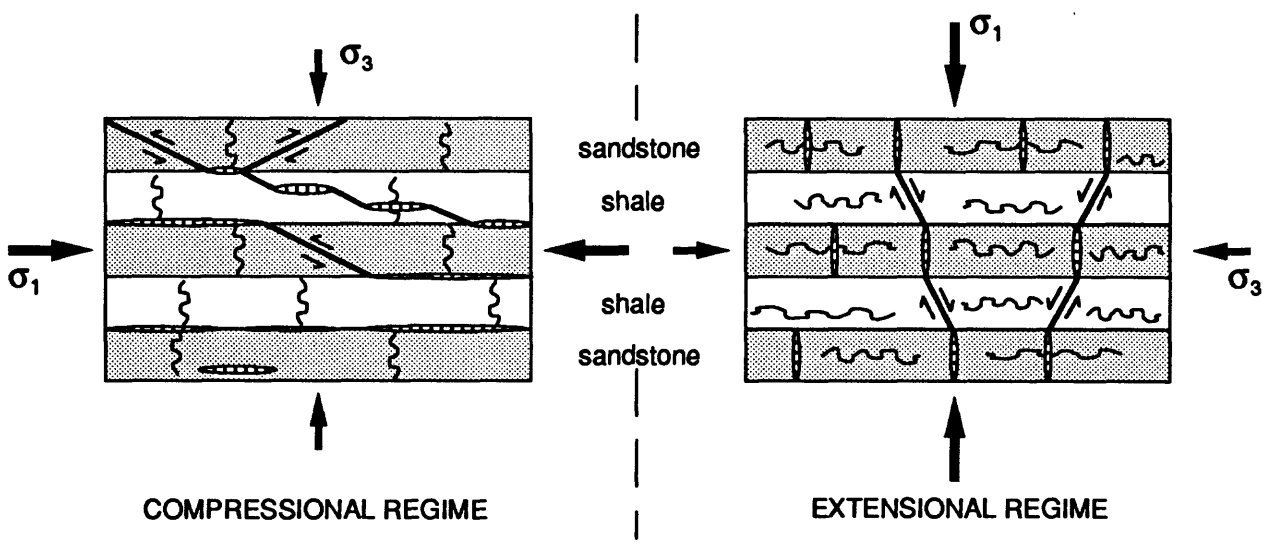



Figure 2 - Schematic representation of stress-control of permeability by "brittle" structures (faults, extension fractures (hachured), and stylolites (squiggly lines)) affecting an anisotropic, sandstone-shale sequence in compressional and extensional stress regimes.

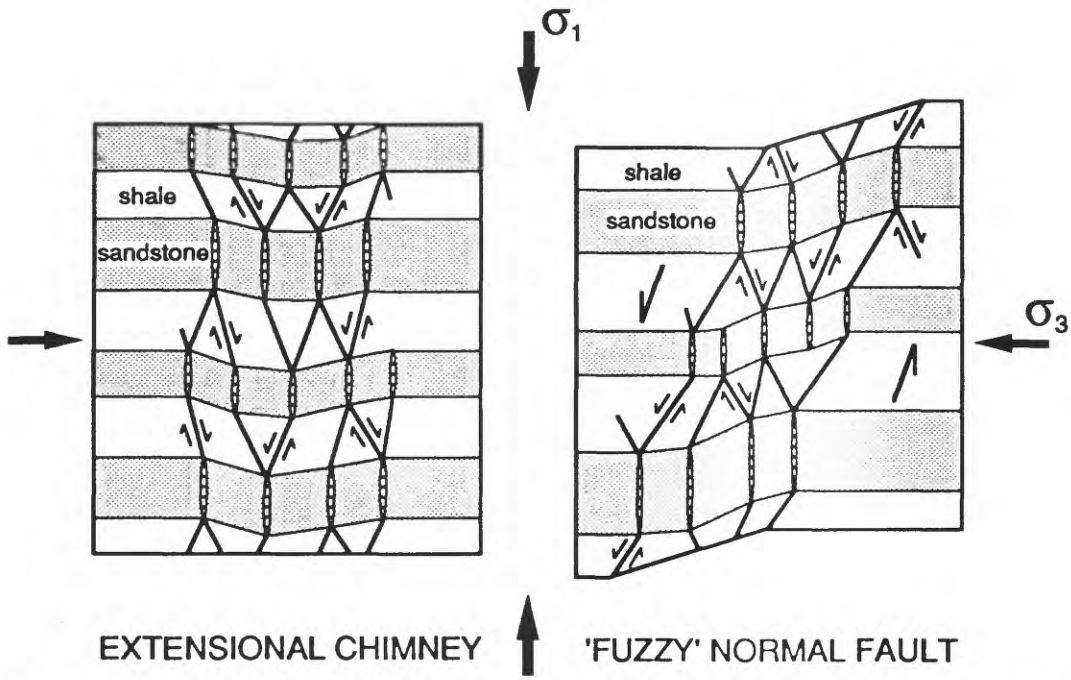



Figure 3 - Hill fault/fracture meshes acting as fluid conduits within an extensional stress regime, either as extensional chimneys or as 'fuzzy' normal faults. Note that the passage of large fluid volumes would lead to increased disorganisation and brecciation within the fault/fracture mesh.

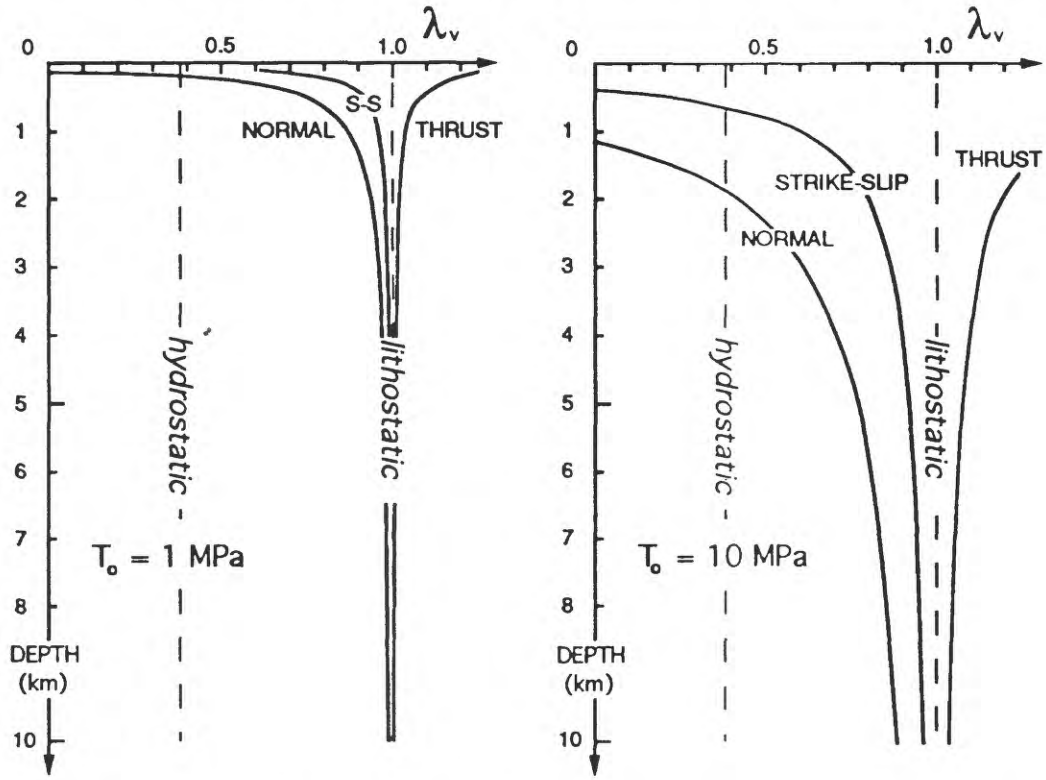



Figure 4 - Fluid pressure conditions for hydraulic extension fracturing versus depth in the stress regimes for Normal, Strike-slip (S-S), and Thrust faulting (after Secor, 1965). Fluid pressure is expressed in terms of the pore-fluid factor ($\lambda_v = P_f/\sigma_v$). The plots are drawn for rock masses with tensile strengths of 1 MPa and 10 MPa, bracketing the common range for sedimentary rocks.

Figure 5 - Normal-fault/fracture mesh incorporating an extensional breccia chimney with infillings of calcite and bitumen from the Monterey Formation, Arroyo Burro Beach, Santa Barbara.

# FRACTURING IN NORMAL FAULT ZONES: IMPLICATIONS FOR FLUID TRANSPORT AND FAULT STABILITY

Ronald L. Bruhn
Department of Geology and Geophysics
University of Utah
Salt Lake City, Utah 84112

## Abstract

Field observations and measurements of fracturing provide important constraints on structural and mechanical properties of large normal fault zones. Large fault surfaces that rupture during earthquakes are encased in intensely fractured rock. The fracture networks provide pathways for fluid flow within the fault zone, and form fluid reservoirs. Estimates of fracture permeability vary from a high of $10^{-13}$ $m^2$ in a pre-rupture stress state marked by high fluid pressure, to a low of $10^{-17}$ $m^2$ in a post-rupture state with hydrostatic fluid pressure. High pressure fluid trapped in fractured-rock reservoirs may bleed into voids between fault surfaces, possibly triggering instability by the reduction of effective normal stress.

## Introduction

Fault zones are characterized by intense fracturing and hydrothermal alteration. Displacement is localized in a slip-zone of cataclasite, breccia and phyllonite surrounding corrugated and striated fault surfaces (Fig. 1). Slip-zone rock grades into fractured, but less comminuted and hydrothermally altered rock in the transition-zone, which grades abruptly into the wall rock. Permeability is dominated by fracturing but is time-dependent because of fracture sealing and filling by precipitated minerals (Bruhn et al., 1990; Blandpied et al., 1992).

Hydrologic properties of fault zones are fundamental to problems in earthquake mechanics, mineral deposition, hydrocarbon production, and ground water transport (Sibson, 1989; Bruhn et al., 1990; Smith et al., 1990; Byerlee, 1990,1993; Blandpied et al., 1992; Rice, 1992; Boutellier and Robert, 1992). Characterizing fracture networks is difficult because of inaccessibility, limited core retrieval, and partial destruction or overprinting of fracture networks during uplift and exhumation of fault zones. However, advancing our

understanding of fault mechanics requires more knowledge of the structural, geophysical, and fluid transport properties of fracture networks.

This report describes a reconnaissance survey of fracturing in the Dixie Valley fault zone, a large normal fault in the Central Nevada Seismic Belt. The spatial distribution, geometry, and intensity of fracturing is described, and fracture permeability is estimated for stress states that may be representative of pre-rupturing and post-rupturing regimes in large normal faults. Finally, we present a preliminary analysis that treats fracture networks as potential reservoirs of high pressure fluid, and consider the effects of injecting this fluid between fault surfaces during the onset of rupturing.

**Dixie Valley Fault Zone**

The Dixie Valley fault is located in the Central Nevada Seismic Belt, and last ruptured during a M = 6.8 earthquake in December, 1954 (Slemmons, 1957). The fault cuts and deforms a 25 to 29 Ma old granitic body and overlying Tertiary volcanic rocks along the eastern margin of the Stillwater Mountains (Speed and Armstrong, 1971; Wilden and Speed, 1974). Vertical displacement is estimated as 6 km (Okaya and Thompson, 1985; Parry et al., 1991). Fault surfaces dip 30° to 45° ESE on average, and slickenlines and grooves trend 105° ± 7°. The fault zone is marked by a band of hydrothermal alteration and intense fracturing that is several ten's to hundred's of meters wide and extends for ten's of kilometers in the footwall adjacent to Quaternary scarps (Fig. 2). Locally, the alteration and fracturing extends hundreds of meters into the footwall along narrow zones of intense cross fracturing (Parry et al., 1991).

The slip-zone consists of breccia and cataclasite, together with corrugated fault surfaces (Fig. 1) that are preserved at only a few sites within the area shown in Fig. 2. Transition-zone or 'damage-zone' rock is intensely, but heterogeneously fractured (Fig. 1). Fracturing is characterized by three prominent sets (Fig. 3A): (1) Shear fractures, most of which dip towards the ESE, but some dip WNW. (2) Steeply dipping fractures that strike ENE, parallel to the trend of the fault zone. These fractures developed mostly by extension, but many have undergone some shearing. (3) Cross fractures that strike at moderate to high angle to the trend of the fault zone, dip at steep to moderate angle, and form a mixed set of both extensional and sheared fractures. Many fractures in these three sets have undergone multiple stages of hydrothermal alteration, and contain cemented cataclastic grains that were derived by abrasion during fault slip, or injected as a cataclastic

slurry during faulting and fluid pressure fluctuations (Parry et al, 1991). Parry et al. (1991) conclude that the hydrothermal mineral assemblages formed at temperatures as high as 305° C and fluid pressure up to 157 MPa, corresponding to depths of 6 km and nearly lithostatic fluid pressure. Alteration continued during uplift of the footwall by faulting, so that lower temperature and pressure alteration minerals were superimposed on the deeper, higher T and P mineral assemblages. The most important mineral assemblages, in rank from highest to lowest T and P, are (1) biotite - K-feldspar, (2) chlorite + epidote, (3) sericite + kaolinite + smectite, and (4) zeolite + clay.

## Field and Computational Techniques

### *Field Measurements of Fracturing*

Fracture intensity and orientation were measured at several sites along the length of the Dixie Valley fault zone (Fig. 2). Fracture intensity was measured by counting the number of fractures intersecting a scan-line over a 10 m interval, and averaged at one meter intervals. Several scan-lines were laid out at each site in order to cross fractures at high angle to strike. Photographs were taken of appropriately oriented rock faces at several sites, fracture traces were marked on the photos, and the trace lengths measured (Fig. 4).

### *Estimates of Permeability*

Fluid transport properties of the fractured rock mass are estimated using the model and algorithms of Oda et al. (1987). Permeability tensor **K** is defined by

$$\mathbf{K}_{ij} = \lambda(\mathbf{P}_{kk}\delta_{ij} - \mathbf{P}_{ij}) \qquad (1)$$

$$\mathbf{P}_{ij} = (\Pi/\ 4{\cdot}V)\ \Sigma_{k=1,N}\,(L_k{}^2[t_k{}^3]\,n_{i\,k}\ n_{j\,k}) \qquad (2)$$

Fractures are modeled as discs of diameter L. $L_k$ is the diameter and $t_k$ is the aperture of the $k^{th}$ fracture in the model network. $n_{ik}$ are the three components of the direction cosine of the $k^{th}$ fracture. $\delta_{ij}$ is the Kroenecker delta function, and V is the volume of the fractured rock mass. N is the number of fractures in the network. $\lambda$ varies from 1/12 for flow between parallel plates of infinite extent to zero for a non-percolating fracture network. $\lambda$ is defined as 1/16 (0.064) for calculations in this study, based on the discussion and numerical simulations of Oda et al. (1987).

Fracture aperture (t) is difficult to assign with any degree of confidence. Aperture decreases rapidly with increasing effective normal stress, and is also a function of the surface roughness and fracture size (Oda et al., 1987). However, none of these functional relationships are well established for fractures at elevated temperature and pressure in natural fault zones. In this study, the initial or 'zero effective stress' aperture ($t_o$) of each fracture is proportional to the fracture diameter, using the ratio ($t_o$/L) = $1x10^{-3}$. The decrease in aperture caused by effective normal stress is computed using the algorithm and coefficients proposed by Oda (1986).

Implementation of equations (1) and (2) requires simulation of a fracture network based on field measurements of fracture intensity, orientation and trace length. The number of fractures in a rock volume is estimated from the scan-line measurements (Fig. 2). The density of fracture poles expressed as % per 1% area on a lower hemisphere stereoplot (Fig. 3A) provides the probability density distribution from which we generate a sample of fractures with the appropriate angular dispersion. Trace length measurements are corrected for sampling bias (Warburton, 1980) and used to assign mean diameters to fractures falling within the orientation ranges of the three prominent fracture sets. Fracture diameters are generated with a negative exponential distribution and the mean diameter specified for each fracture set in Table 1.

The above procedure is automated in a computer program. The code generates an interpolation table of fracture pole orientation density based on field measurements. The operator specifies the total number of fractures to be modeled within a volume of rock, and the mean diameter of fractures in various sets defined by restricted ranges in orientation. A list of fractures is created with the appropriate distribution in orientation and diameter. The fracture list (fracture network) is entered into another computer program which implements equations (1) and (2). This latter code computes the initial aperture ($t_o$) of each fracture as a function of diameter, corrects $t_o$ for closure caused by the specified stress state, and determines the principal magnitudes and directions of the permeability tensor (Table 2, Fig. 3B). Oda et al. (1987) provide a thorough discussion and numerical tests of this 'equivalent porous media' algorithm.

## Results

*Fracture Intensity*

Mesoscopic fracture intensity varies markedly along the length of the fault zone, but fracturing is concentrated in a band up to 300 m wide immediately adjacent to the Quaternary scarps (Fig. 2). Steeply dipping extension fractures strike parallel to the fault zone and occur both as short pinnate fractures originating from fault surfaces (Fig. 4) and as long (meter scale) fractures that truncate several fault surfaces in vertical succession. The scan-line intensity of this latter fracture set is highly variable, but locally exceeds 30/meter. Shear fractures, or fault surfaces, mostly dip ESE and intensity varies from 4/m to 10/m. Cross fracture intensity ranges from less than 4/m to almost 30/m, and locally these fractures coalesce into elongate bands of intense fracturing with hydrothermal alteration that extends several hundred meters into the footwall. The cross fractures originate by at least two different processes, 1) local stress concentrations in the wall rock created by spatial variability in slip along corrugated shear surfaces (Fig. 1), and 2) large scale flexing of the footwall during uplift. Microfracturing mimics the mesoscopic fracturing. Thin section study reveals that microfractures are healed and filled with alteration minerals as the result of the same rock-fluid reactions that occurred along mesoscopic fracture surfaces (Parry et al., 1991).

*Fracture Network Model*

A 'generic' fracture network model is created using representative fracture intensity measurements, orientation, and trace length data (Figs. 2, 3A, 4). The general characteristics of the model are summarized in Table 1, for 1 $m^3$ rock volume. This model is used in subsequent calculations to estimate fracture network permeability and fluid reservoir characteristics. Fracture network heterogeneity is not considered in this preliminary study.

The number of fractures per unit volume ($\rho$) is estimated from scanline intensity values (eqn. 25, Oda et al., 1987).

$$\rho\langle L^2 \rangle = 4I_S / [\pi \langle | \mathbf{n} \cdot \mathbf{q} | \rangle] \qquad (3)$$

$\langle L^2 \rangle$ is the mean squared fractured diameter, $I_S$ is scanline intensity (#/m), and $\langle | \mathbf{n} \cdot \mathbf{q} | \rangle$ is the mean value of the dot product between the poles to fractures in a specific set (defined by unit vectors **n**) and the unit vector parallel to the scanline (**q**). For example, consider the histograms of scanline intensity for the three fracture sets in Figure 2. Using Figure 2, we assign $I_S$ = 8/m for both extension and cross fractures (set 1) and (set 3), and

5/m for shear fractures (set 2). These are average intensities for each fracture set (Fig. 2). Trace length measurements for each set yield in sequence, $\langle L \rangle$ = 0.2 m (set 1), 0.5 m (set 2) and 0.25 m (set 3) (Table 1). Fracture density is estimated from (3) with $\langle | \mathbf{n} \cdot \mathbf{q} | \rangle$ = 1.0. $\rho_1 = 255/m^3$, $\rho_2 = 26/m^3$, $\rho_3 = 164/m^3$, for a total mesoscopic fracture density of 445 fractures/$m^3$. Field and microscopic studies indicate multiple episodes of fracture sealing and filling by hydrothermal minerals. We account for the possible effects of hydrothermal alteration and cementation by assuming that only 1/3 of these fractures are open and transport fluid at any given time. This proportion of open fractures is only a guess, and is not based on any quantitative information. The real value may be significantly higher, or much lower.

The fluid transport and reservoir properties of fault zone fracture networks are important in theories of rupture initiation and modeling of earthquake precursors (Sibson, 1989; Byerlee, 1993). Most large, normal faulting earthquakes initiate at depth ≥ 10 km. For purposes of modeling we impose stress conditions and fluid pressures that may exist in parts of a normal fault zone at a depth of 10 km, based on studies of fluid inclusions and mineral alteration assemblages (Parry and Bruhn, 1990; Bruhn et al., 1990). In the pre-rupture state, the ratio of fluid pressure (Pf) to lithostatic pressure (PL) is defined as $P_f/P_L$ = 0.9. The vertical principal stress ($S_V$) = $\rho_r g z(1-P_f/P_L)$ = 27 MPa, for $\rho_r$ = 2700 Kg/$m^3$. The major fault surfaces have a coefficient of friction $\mu$ = 0.6, no cohesion, and dip 45°. The least principal stress $S_{hmin}$ = 6.8 MPa at failure. For simplicity, we assume that $S_{hmax} = S_{hmin}$, or

$$\phi_s = [S_{hmax} - S_{hmin}]/[S_V - S_{hmin}] = 0.0 \qquad (4)$$

Fluid pressure presumably drops within the transition-zone either during or after rupturing because permeability increases, and newly formed fractures provide connections to the surface (Sibson, 1989; Parry and Bruhn, 1990; Bruhn et al., 1990). Assume that the post-rupture fluid pressure becomes hydrostatic, that the shear stress is completely relaxed, and the three principal stresses become equal, with a magnitude of 170 MPa. These pre- and post-rupture stress states are chosen to provide upper bounds on fracture volume changes and fluid production from the transition-zone at seismogenic depths.

## Discussion

Partial closure of fractures between pre- and post-rupture stress states has a marked effect on fracture permeability (Fig. 3B, Table 2). The maximum and intermediate permeability axes (**K**1 and **K**2) plunge more steeply than the plane of the fault zone, and the least principal axis (**K**3) plunges more gently than the pole to the fault zone in the pre-rupture stress state (Fig. 3B). The average fracture aperture is 19 microns, and total fracture volume is 0.52 liter. Fracture permeability is large in the pre-rupture state due to low effective normal stress across most fractures. Permeability magnitudes are $\approx 10^{-13}$ $m^2$, about 4 orders of magnitude greater than in the post-rupturing stress state (Table 2).

A marked reduction in permeability in the post-rupture stress state is accompanied by a rotation in permeability axes. **K**1 and **K**2 are located essentially within the plane of the fault zone, and **K**3 is almost parallel to the fault zone pole (Fig. 3B, Table 2). Average aperture is 1.5 microns, and fracture volume is 0.04 liter. Permeability magnitude is $\approx 10^{-17}$ $m^2$.

The increase in effective stress between pre- and post-rupture stress states could produce about 0.5 liter of fluid from a cubic meter of fractured, transition-zone rock because of partial closure of fracture apertures. If fluid expulsion occurs during the initial phase of rupturing, this fluid may be injected between fault surfaces and act to trigger an earthquake (Byerlee, 1993). The fluid also acts as a mineralizing agent, carrying ionic species in solution into fractures and voids between fault surfaces, where chemical reactions produce cement and mineral filling that ultimately reduces permeability, and partially heals fault surfaces.

Byerlee (1993) proposes a 'fluid compartment' model for fault zones and the triggering of earthquakes. Fluid sealed under high pressure in one compartment is partially drained into an adjacent compartment of initially lower fluid pressure. We use this model to investigate the potential effect of draining fluid from the transition-zone fracture network into the space between opposing walls of faults or 'slip-surfaces' (Fig. 1). Assume the following initial conditions at a depth of 10 km: 1) Fluid is initially trapped at hydrostatic pressure in the fault, and 2) fluid is trapped at 0.9 $P_L$ in the adjacent fracture network. The fault and fracture network are hydraulically isolated until the onset of rupturing, when the seal between the two compartments is broken. Fault permeability is much less than that of the fracture network, and we neglect changes in fluid compressibility during pressure changes.

The final fluid pressure, after equalization of pressure between the two compartments is (Byerlee, 1993):

$$P_f = [V_{tz}P_{tz} + V_{sz}P_{sz}]/[V_{tz} + V_{sz}] \qquad (5)$$

V and P are the fracture volumes and initial fluid pressures in the transition-zone (tz) and fault (sz), respectively. We assign $V_{tz} = 0.52 \times 10^{-3}$ m$^3$, and $P_{tz} = 0.9$ lithostatic pressure ($P_{tz} = 243$ MPa) based on the stress and transition-zone structural parameters summarized in Tables 1 and 2.

How large a section of fault surface is affected by injection of fluid from 1 m$^3$ of transition-zone rock? Consider the situation where $P_f = 171.5$ MPa, half the difference between $P_{tz}$ and $P_{sz}$. Rearranging equation 5 to solve for $V_{sz}$ gives:

$$V_{sz} = V_{tz}[P_{tz} + P_f]/[P_f - P_{sz}] \qquad (6)$$

The fault area that can be pressurized from $P_{sz} = 100$ MPa to $P_f = 171.5$ MPa is $V_{sz}/t_{sz}$, where $t_{sz}$ is the average aperture between opposing walls of the fault. $t_{sz} \approx 1$ to 5 microns for the specified stress conditions according to the joint aperture constitutive law of Oda (1986). A fault surface area $\geq 100$ m$^2$ is affected by the fluid injected from 1 m$^3$ of transition-zone rock. This preliminary result is intriguing because fluid from only a small volume of over-pressured rock may apparently reduce the effective normal stress over a large fault surface area. Only small, heterogeneously distributed patches of over-pressured fracture network are required to make the fluid pressure compartment model viable. This conclusion is consistent with observations that the structural and fluid pressure properties of fractured rock are spatially and temporally heterogeneous (Bruhn et al., 1990; Parry and Bruhn, 1990).

## Conclusions

Field observations and measurements of fracture networks surrounding large fault surfaces provide important constraints on structural and mechanical properties of large normal fault zones. Large fault surfaces that rupture during earthquakes are encased within intensely fractured rock. The fracture networks provide pathways for fluid flow within the fault zone, and also form fluid reservoirs. High pressure fluid trapped in isolated, fractured-

rock reservoirs may bleed into voids between fault surfaces, possibly triggering instability by the reduction of effective normal stress, as envisaged by Byerlee (1993). Estimates of fracture permeability vary from as high as $10^{-13}$ $m^2$ in the modeled pre-rupture stress state, to as low as $10^{-17}$ $m^2$ in the post-rupture state.

**Acknowledgements**

This work was partially supported by grants from the National Science Foundation awarded to R.L. Bruhn and W.T. Parry, and by the Nationl Earthquake Hazards Reduction Program of the United States Geological Survey, awarded to C.B. Forster and R.L. Bruhn..

**References**

Blanpied, M.L., Lockner, D.A., and Byerlee, J.D., 1992, An earthquake mechanism based on rapid sealing of faults. Nature, V. 358, p. 574-576.

Boutellier, A.-M., and Robert, F., 1992, Paleoseismic events recorded in Archean gold-quartz vein networks, Val d'Or, Abitibi, Quebec. J. Struct. Geol., V. 14, p. 161-179.

Bruhn, R.L., Yonkee, W. T., and Parry, W. T., 1990, Structural and fluid-chemical properties of seismogenic normal faults. Tectonophysics, V. 175, p.139-157.

Byerlee, J.D., 1993, Model for episodic flow of high-pressure $H_2O$ in fault zones before earthquakes. Geology, V. 21, p. 289-304.

Byerlee, J.D., 1990, Friction, overpressure and fault normal compression. Geophys. Res. Letts., V. 17, p. 2109 - 2112.

Oda, M., Hatsuyama, Y., and Ohnishi, Y., 1987, Numerical experiments on permeability tensor and its application to jointed granite at the Stripa Mine, Sweden. J. Geophys. Res., V. 92, p. 8037-8048.

Oda, M., 1986, An equivalent continuum model for coupled stress and fluid flow analysis in jointed rock masses. Water Resources Res., V. 22, p. 1845-1856.

Okaya, D.A., and Thompson, G.A., 1985, Geometry of Cenozoic extensional faulting, Dixie Valley, Nevada. Tectonics, V. 4, p. 107-125.

Parry, W.T., Hedderly-Smith, D., and Bruhn, R.L., 1991, Fluid inclusions and hydrothermal alteration on the Dixie Valley Fault, Nevada: J. Geophys. Res., v. 96, p. 19,733-19,748.

Parry, W. T., and Bruhn, R. L., 1990, Fluid pressure transients on seismogenic normal faults: Tectonophysics, v. 179, p. 335-344.

Rice, J.R., 1992, Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault: *in* Fault Mechanics and Transport Properties of Rocks, B. Evans and T. -F. Wong (eds)., p. 475-504 (Academic Press).

Sibson, R.H., 1989, Earthquake faulting as a structural process. J. Struct. Geol., V. 11, p. 1-14.

Slemmons, D.B., 1957, Geological effects of the Dixie Valley-Fairview Peak, Nevada, earthquakes of December 16, 1954. Bull. Seism. Soc. Amer., V. 47, p. 353-375.

Smith, L., Forster, C.B., and Evans, J.P., 1990, Interaction of fault zones, fluid flow and heat transfer at the basin scale: *in* Hydrogeology of permeability environments. Inter. Assoc. of Hydrogeologists, V. 2, p. 41-67.

Speed, R.C., and Armstrong, R.L., 1971, Potassium-Argon ages of some minerals from igneous rocks of western Nevada. Isochron/West, V. 71-1, p. 1-8.

Warburton, P.M., 1980, A stereological interpretation of joint trace data. Int. J. Rock Mech. Min. Sci. & Geomech. Abstr., V. 17, p. 181-190.

Wilden, R., and Speed, R.C., 1974, Geology and mineral resources of Churchill County, Nevada. Nevada Bureau of Mines and Geology Bulletin 83, 95 p.

## TABLE 1: FRACTURE SET PARAMETERS

| Fracture Type | Pole Direction (°) Trend | Plunge | Mean Trace Length (m) | Fracture Density* ($\rho$ = #/$m^3$) |
|---|---|---|---|---|
| **Extension** | | | | 255 |
| Set #1a | 260 - 320 | 0 - 29 | 0.2 | |
| Set #1b | 105 - 141 | 0 - 30 | 0.2 | |
| **Shear Fractures** | | | | 26 |
| Set # 2 | 250 - 315 | 30 - 65 | 0.5 | |
| **Cross Fractures** | | | | 164 |
| Set #3a | 180 -245 | 30 - 70 | 0.25 | |
| Set #3b | 003 -050 | 40 - 75 | 0.25 | |

* Computed from scan-line intensity for each fracture set. Total # of mesoscopic fractures in1 $m^3$ of rock is 445.

## TABLE 2: ESTIMATED FRACTURE PERMEABILITY

| Principal Stresses (MPa) | | | | Permeability ( $10^{-13}$ $m^2$) | | | Fracture Vol. (liters) |
|---|---|---|---|---|---|---|---|
| Case | $s_V$ | $S_{hmax}$ | $S_{hmin}$ | K1 | K2 | K3 | |
| 1 | 27 | *6.8 @ 020° | 6.8 | 1.2 !(023/38) | 1.1 (178/49) | 0.3 (283/13) | 0.52 |

| Principal Stresses (MPa) | | | | Permeability ( $10^{-17}$ $m^2$) | | | Fracture Vol. (liters) |
|---|---|---|---|---|---|---|---|
| Case | $s_V$ | $S_{hmax}$ | $S_{hmin}$ | K1 | K2 | K3 | |
| 2 | 170 | *170@ 020° | 170 | 5.0 !(032/23) | 4.7 (142/38) | 1.3 (278/43) | 0.04 |

Case 1: Stress state just before rupturing.
Case 2: Stress state shortly after rupturing.

! Trend and plunge of permeability tensor axes.

$S_V$ = vertical principal stress, $S_{hmax}$ = maximum horizontal stress, $S_{hmin}$ = minimum horizontal stress.

* Maximum Horizontal Stress trends 020°, parallel to fault zone.

$\phi s = [S_{hmax} - S_{hmin}]/[S_V - S_{hmin}]$

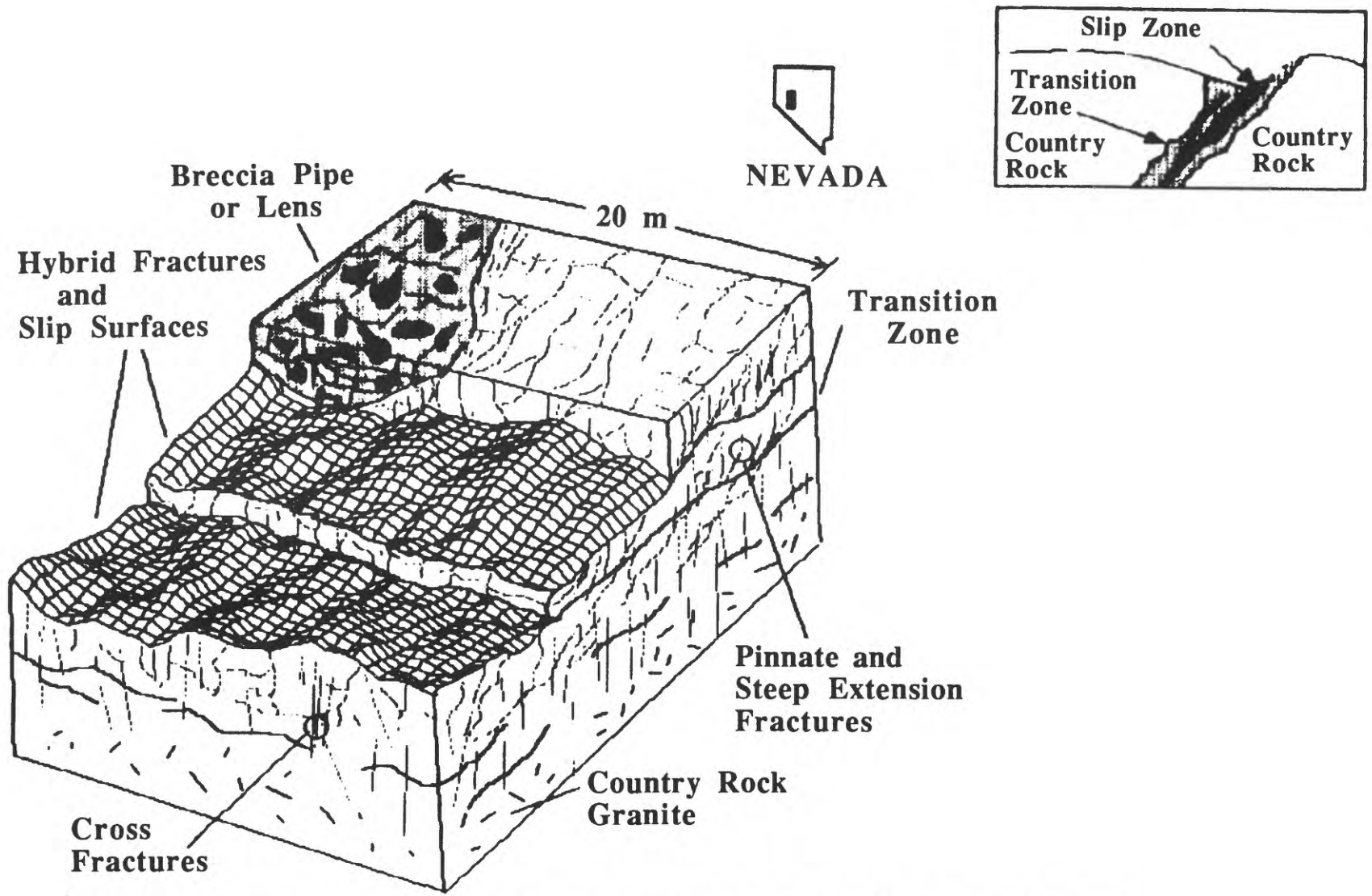



Figure 1: Diagram of corrugated fault surfaces and transition zone fracture networks based on observations in the Dixie Valley fault zone. Inset shows position of slip-zone and transition zone rock within the fault zone.

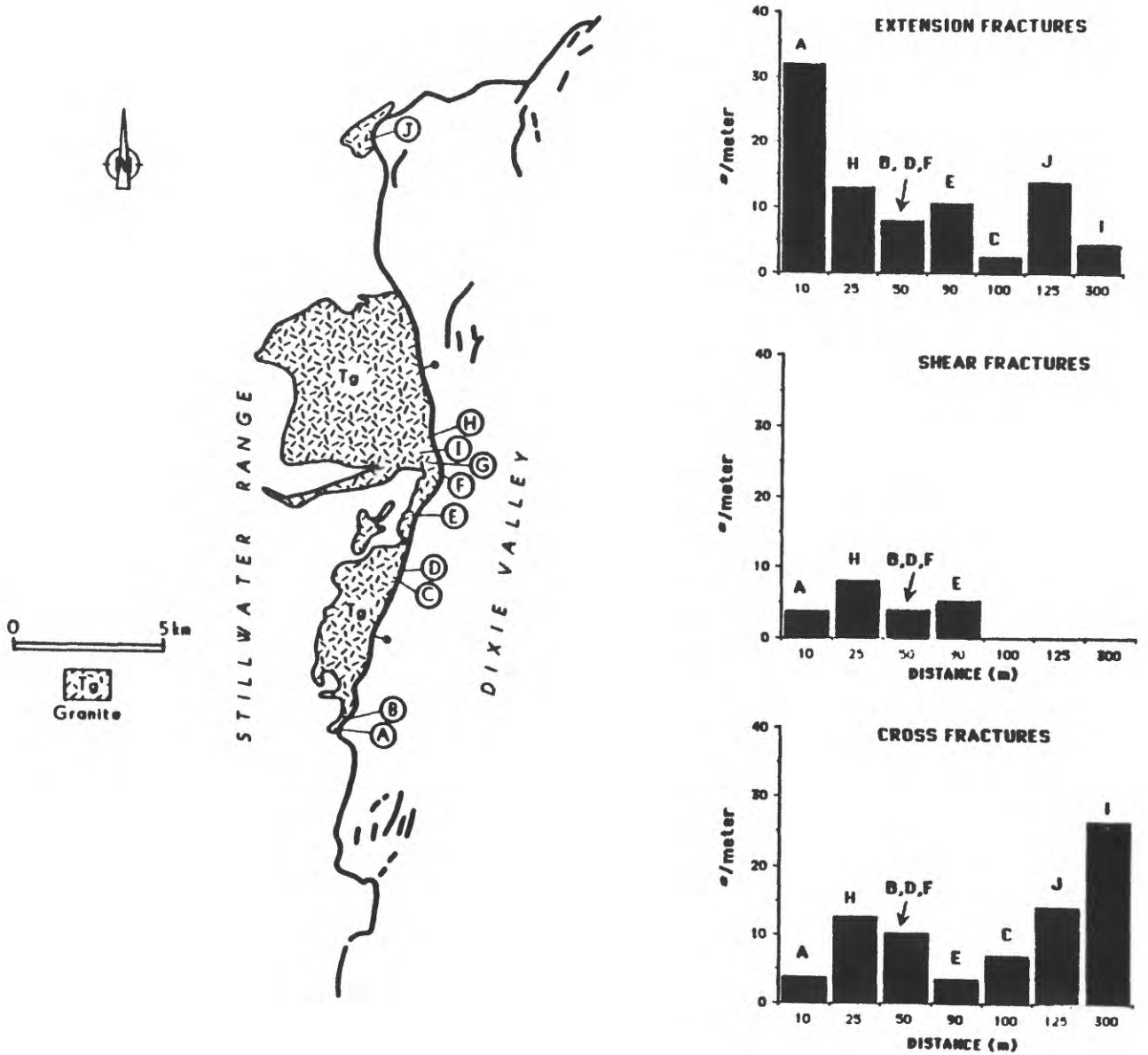



Figure 2: Spatial variations in fracture intensity along the Dixie Valley, Nevada fault zone. Location of each measurement site is marked on the geologic map, and coded on the bar graphs. Refer to Figure 1 and the text for discussion of the three prominent fracture types.

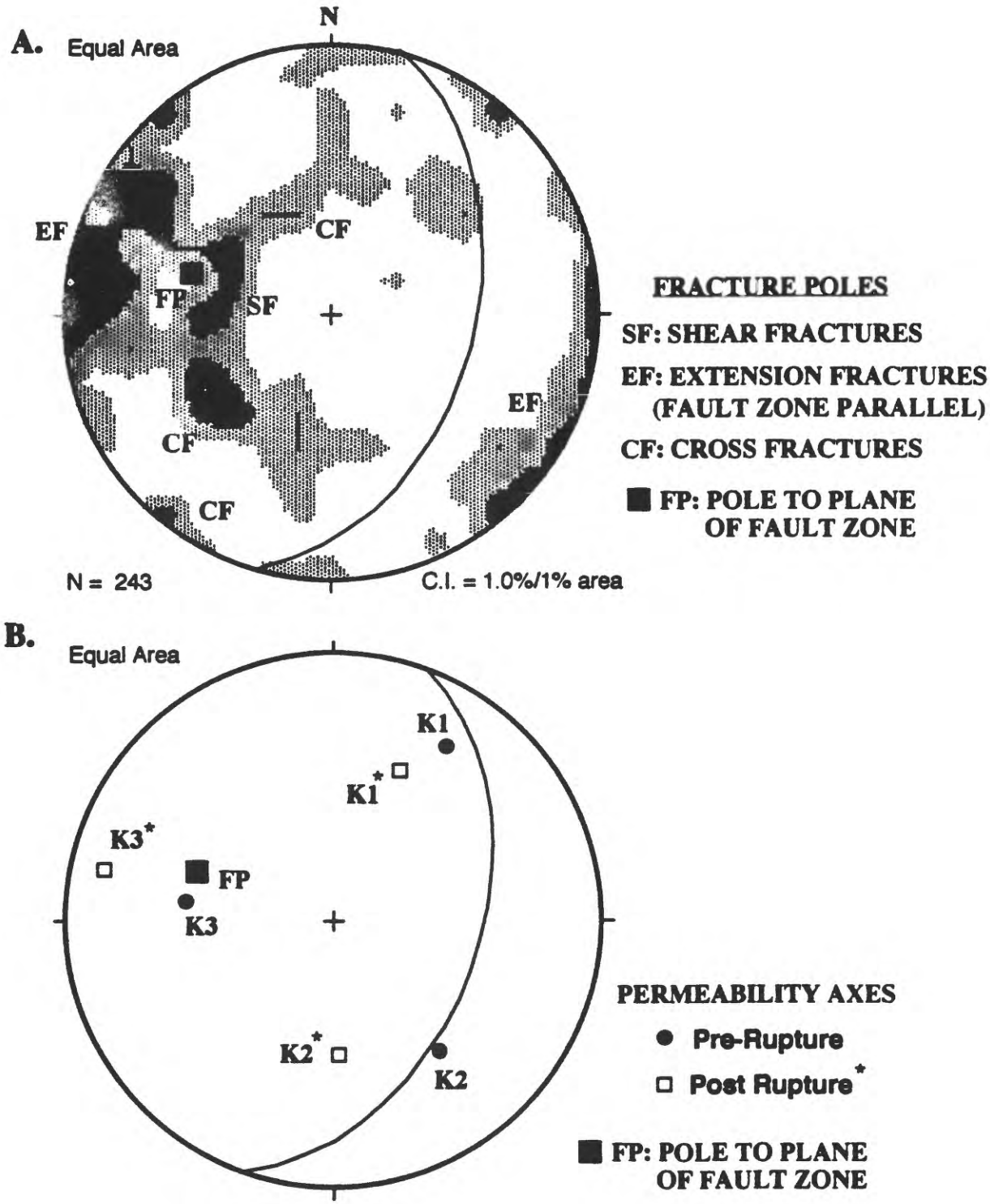



Figure 3: A. Contoured stereoplot of poles to fractures in the Dixie Valley fault zone. Contoured by the Schmidt method in intervals of 1%/1 % area of a lower hemisphere. The regions associated with poles to shear, extension and cross fractures are marked on the diagram. The great circle trace represents the average plane of the fault zone, and the pole (FP) is perpendicular to the fault zone.
B. Stereoplot of permeability axes for pre-ruputure and post-rupture stress states. The pole (FP) and great circle trace represent the average fault zone orientations.

## FRACTURE NETWORK: VIEW ALONG FAULT STRIKE

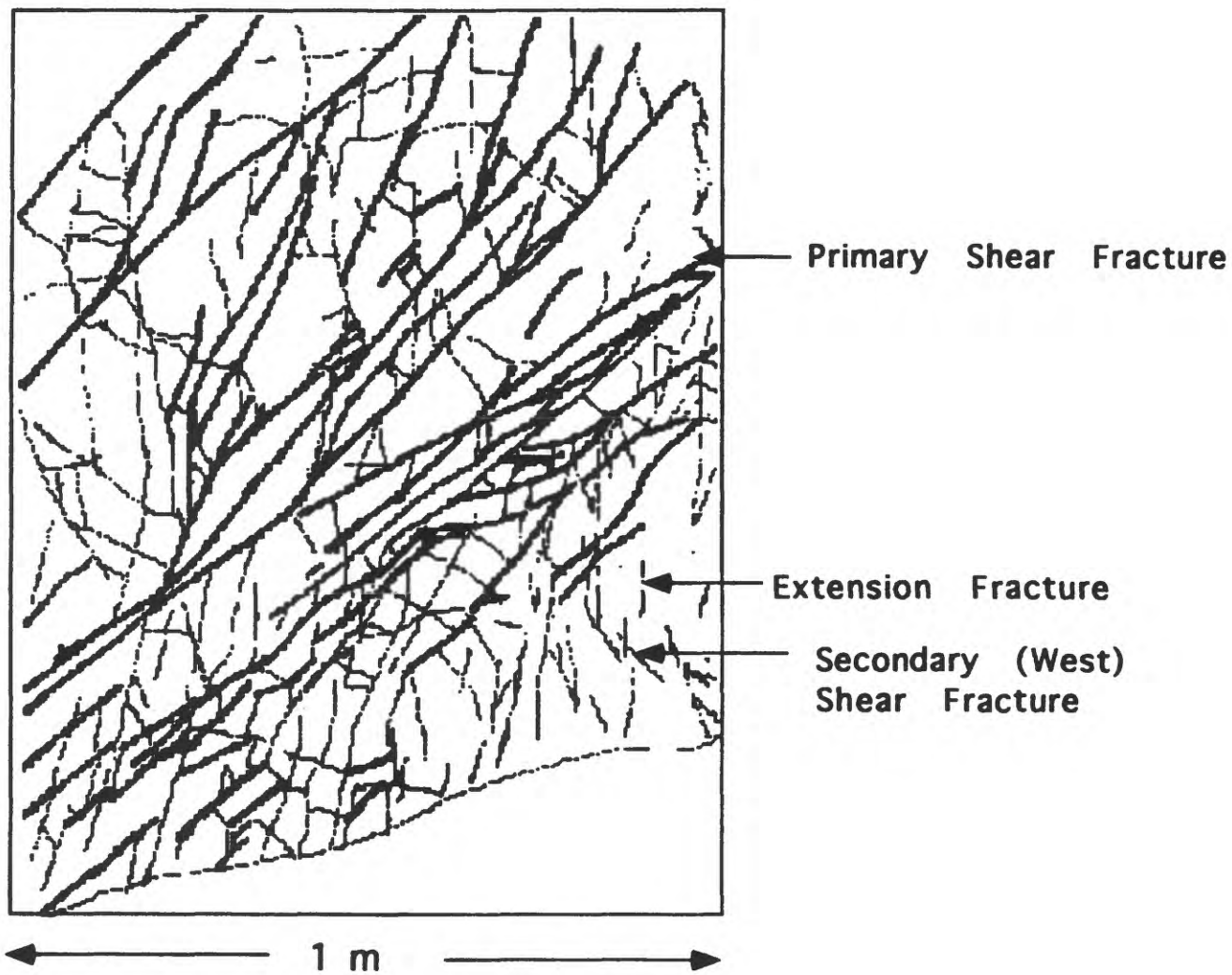



Figure 4: Fractures traced from a photograph of a rock face in the Dixie Valley fault zone. The view is along strike, and the rock face is perpendicular to the strike of the fault zone.

# The Permeability of Faults

C.H. Scholz and M.H. Anders
Lamont-Doherty Earth Observatory and Dept. of Geologic Sciences
Columbia Univ., Palisades, N.Y., 10964

## Abstract

Faults are observed to play a profound role in fluid migration in the crust. Here we discuss, in generic terms, the structures within fault zones that influence and control the hydraulic transport properties of the fault. There are two such structures, the cataclasite zone (ccz), the zone of crushed rock within the fault core, and the process zone wake (pzw), the zone on enhanced cracking in the wall rock just adjacent to the fault core. The ccz is produced by wear processes due to cumulative fault slip, and hence scales in width with fault displacement. The process zone wake is produced by within the stress concentration at the fault tip and hence is created when the fault tip propagates past the point in question during fault growth. The width of this zone scales progressively with fault length. These structures may subsequently be sealed by mineralization or clay minerals, which need not affect both structures simultaneously. Hence for general possibilities result, a) ccz and pzw both open, fault acts as both a transverse and longitudinal conduit; b) ccz sealed, pzw open, fault is a barrier to transverse flow and as an outer conduit to longitudinal flow; c) ccz open, pzw sealed, fault acts as a tranverse barrier and as an inner conduit to longitudinal flow; d) ccz and pzw both sealed, fault acts as a general barrier.

## Introduction.

In many situations of broad practical interest fluid flow through the uppermost crust of the earth is controlled by fractures, which in consolidated rock are found ubiquitously on all scales. The usual approach in such studies is to characterized the fracture as an irregularly open planar conduit and to consider flow through one or a network of such channels. This is a good description of joints, natural tensile fractures, which in most places are the most plentiful fractures. However, though less plentiful, the largest fractures are often faults, and faults, owing both to their mode of origin as shear fractures and because of their accumulated shear displacements, have very different properties than joints, particularly with respect to fluid transport. It is well known both in oil field and mining experience that faults may sometimes act as fluid conduits and in other cases as impermeable barriers. This presents a severe problem, particularly in petroleum applications, because although faults can be readily imaged with geophysical remote sensing, one cannot tell from the image whether the fault is an aquifer or an aquaclude. In the one case the fault may act as a trap, in the other a migration conduit, and both exploration and reservoir management strategy depend on this distinction being made.

Here we envision faults as *permeability structures.* We begin with a generic model of a fault, described in the next section, in which the several structural elements, each with their permeability characteristics, are linked to the evolution of the fault. There is enough data at present to show that this model is correct in the essentials. We then discuss the overall permeability structure that results.

## A Generic Fault Model.

The point of departure for this paper is a generic model of faults that we have been developing over the past several years, in our efforts to understand the growth of faults. The model is still sketchy in many respects, but enough key data has been collected and we have a firm enough grasp of the mechanics behind it that we are confident that its major

features are broadly correct. This model, and the permeability model that can be derived from it, provide a working hypothesis for testing various aspects of fault-fluid interactions.

The basic elements of the model are shown in Fig. 1. The top frame shows a typical profile of the distribution of slip from the center to the tip of a fault. Slip decreases from a maximum, $D_{max}$ , near the center of the fault and decreases towards the tip. The tapering form of this decrease near the fault tip contrasts with the elliptical slip distribution expected from a simple elastic crack model. This finite slip gradient at the tip results in a near tip stress concentration that is finite and does not contain the singularity of the elastic crack model.

This slip distribution is well modelled with the Dugdale-Barenblatt type of crack model (Cowie and Scholz, 1992a). This is an elastic-plastic model in which there is a breakdown zone S of distributed inelastic deformation in which there is a progressive strength breakdown from intact rock to the fault friction. The predicted stress distribution is shown in the middle frame of Fig. 1, where $\sigma_f$ is the residual frictional stress on the fault, $\sigma_0$ is the applied tectonic stress, and $\sigma_y$ is the yield stress of the rock.

Structures associated with faults are shown in the lower frame of Fig. 1, which is now a section normal to the fault and along its length, rather than a profile along length. There are three structural elements, the brittle process zone, I, which is a zone of distributed cracking in the country rock surrounding the fault, a cataclasite zone II of wear detritus within the fault, and a narrow zone of ultracataclasite or clay fault gouge, III, that may or may not be present within that.

These characteristic fault structures differ greatly from those associated with joints. Although joints may have process zones (Pollard and Segall, 1987), those for faults are much broader and well developed. There are at least two reasons for this. The first is that since a shear crack cannot propagate as such within its plane, a fault grows by first producing a complex mesh of tensile fractures that gradually break down to form a throughgoing fault (Cox and Scholz, 1988). The second is that because rock is about ten times stronger in shear than tension, the stress concentration at the tip of a growing fault will be about ten time that at the tip of a growing joint. This strength difference is well known from laboratory experiments, but can also be demonstrated from field evidence. The ratio of maximum slip to length of a fault, $D_{max}/L$ is approximately equal to $\sigma_y/\mu$, where $\mu$ is the shear modulus of the rock. The equivalent expression for a joint, inflated to the point of propagation, equates this ratio to $A_{max}/L$, where $A_{max}$ is the maximum aperture. Typical observed values of $D_{max}/L$ are of the order of $10^{-2}$ Cowie and Scholz, 1992b), as compared to $A_{max}/L$ values that are typically of order $10^{-3}$ (measured from veins and dikes, Vermilye and Scholz, 1991). Joints, of course, also do not contain cataclasite or ultracataclasite zones, since these are the products of wear during shearing. These are the differences between faults and joints that lead to them being very different in their fluid transport properties.

***Scaling Relationships.*** All of these fault features scale with the size of the fault. The most fundamental of the scaling relationships is between D and L. Although some earlier workers have suggested that D and L scale nonlinearly (Walsh and Watterson, 1988; Marrett and Allmendinger, 1991), we believe the scaling to be linear, as predicted by the Dugdale-Barenblatt model, and these earlier interpretations to have been a result of the limited range and poor quality of the data (Cowie and Scholz, 1992b). We have collected recently a much more homogeneous data set extending over a much broader size range that shows the linearity of this scaling very clearly (Fig. 2, Dawers et al., 1993; Scholz et al., 1993).

The process zone, a zone of enhanced cracking that surrounds the fault, is produced within the stress concentration of the advancing fault tip -- it is a relic of the passage of the

fault tip and reflects the distributed inelastic zone associated with the breakdown region S. The width of the process zone, P, will be defined by the distance to a limiting stress, $\sigma_d$, akin to the "dilatancy stress" in laboratory experiments necessary to produce cracking (Fig. 1, middle frame). Because of the self-similar nature of fault growth, the model predicts that P, like S, scales linearly with fault length, L.

Chernyshev and Dearman (1991) describe observations of process zones for faults in Tajikistan that range in length from 1 to 100 km. They observe it by counting macrofractures, the density of which they find to increase as the fault is approached. Anders and Wilschko (1993), on the other hand, counted the density of microfractures within grains, finding a similar increase in fracture density as the fault is approached. Examples from the latter are shown in Fig. 3. Both find an exponential increase as the fault is approached.

The cataclasite zone is formed by wear due to progressive slip on the fault. It often fines inwards, starting from a coarse breccia and progressing towards a finely milled rock flour or gouge, the nature of the wear material depending fairly strongly on the wall rock from which it is derived. The thickness of this zone, T, is proportional to slip displacement, D (Scholz, 1987). The constant of proportionality depends mainly on the strength and nature of the wall rock and the normal stress, and is typically of the order $10^{-2}$. Within the cataclasite zone are sometimes observed thin lenses or zones of ultracataclasite or clay gouge (III in Fig. 1). These are composed of very fine grained material that often shows evidence of hydrothermal alteration (Chester et al, 1993). These zones are very thin: in cases described by Chester et al. (1993) they are of the order $10^{-3}$T. In what follows, however, we will negect affects due to the this zone.

***Permeability model.*** In the above discussion we established that we can now make some general statements about the structure of fault zones and how those structures scale with such routinely determined measures of fault size such as length or slip. We can now use this mechanical model to construct a permeability model, shown in Fig. 4. What is shown at the top of each figure is permeability in a transect across a fault. The scale is schematic: permeability is a very strong function of crack density and aperture, which we do not attempt to model here. Crack density is enhanced within the process zone wake, pzw, increasing rapidly towards the fault. We consider four possibilities concerning the cataclasite zone (ccz) and process zone wake (pzw). In case 1 (fig. 4a) the cataclasite consists of fragmented, granular material such as what would be derived from nonphylitic and clayfree wall rock. In this case both the ccz and pzw would be a zone of very high permeability, both for flow along and across it. In case 2 (fig. 4b) the cataclasite zone itself is clay-rich, as might be the case if the wall rock were shale or if the material in the ccz wer wheathered, the fault would act as an impermeable barrier to flow across it but still act as a highly permeable conduit for flow along it. We call this the outer conduit case, since flow would occur in the pzw outside the fault core. Finally, the fractures within the fault structure may become sealed by mineralization, in which case two possibilities arise. If the fault is still active, the ccz may be continually refractured, so that it is permeable while the pzw is sealed. This is case 3 (fig. 4c), which we call the inner conduit situation, since flow can only occur within the fault core. The difference between the inner and outer conduit situations may be appreciated by considering a dipping fault. In the outer conduit, fluids migrating upwards will be dammed against the fault and migrate along it, but in the inner conduit situations, only fluid that get into the fault core, say from depth, can migrate along it. Finally, if the fault is inactive, both the ccz and pzw may be sealed by mineralization, resulting in case 4 (fig. 4d), in which the fault becomes an entirely impermeable structure.

So it is no wonder that faults have been observed to play such diverse roles in fluid migration. The problem addressed by this paper is obtain a better understanding of what leads to the final *permeability structure* of a fault, so that criteria can be established

for determining it from external factors such as fault length, slip, depth of burial, and the composition of the wall rocks.

## REFERENCES


Anders, M.H. & Wiltschko, D.V. 1993. Microfracturing, paleostress, and the growth of faults. Subm. J. Struct. Geol.

Chernyshev, S. N. & Dearman, W. R. 1991. *Rock fractures*. Butterworth-Heinemann Ltd, London England 272.

Chester, F.M., R.L. Beigel, and J. Evans, 1993. Internal Structure and weakening mechanism of the San Andreas fault, *J. Geophys. Res. 98,* 771-786.

Cowie, P. A. & Scholz, C. H. 1992a. Physical explanation for displacement-length relationship of faults using a post-yield fracture mechanics model. *J. Struct. Geol.* 14,1133-1148..

Cowie, P. and C. Scholz, C.H. 1992b. Displacement-length scaling relationship for faults: data synthesis and discussion. *J. Struct. Geol.* 14,1149-1156.

Cox, S. J. D. and Scholz, C.H. 1988. On the formation and growth of faults, an experimental study: *J. Struct. Geol.* **10**, 413-430.

Dawers,N., M. Anders, and C. Scholz, 1993. Evidence for self-similar fault growth from the Volcanic Tableland, eastern Calif., in prep.

Marrett, R. L. and Allmendinger, R.W. 1991. Estimates of strain due to brittle faulting: sampling of fault populations. *J. Struct. Geol.* **13**, 735-737.

Pollard, D. D. and Segall, P. 1987. Theoretical displacement and stresses near fractures in rock: with applications to faults, joints, veins, dikes and solution surfaces. **in** Fracture Mechanics of Rock. ed. Atkinson, B.K., Academic Press. 277-347.

Scholz, C. H. 1987. Wear and gouge formation in brittle faulting. *Geology* **15**, 493-495.

Scholz, C.H., N.H. Dawers, J.Z. Yu, M.H. Anders, and P.A. Cowie, 1993. Fault growth and fault scaling laws: preliminary results, *J. Geophys. Res.*, in press.

Vermilye, J. and Scholz, C.H. 1991. Fracture length and aperture, a general relationship. *Proc. 13th Annual PL/DARPA Seimic Research Symposium.*

Walsh, J. J. and J. Watterson. 1988. Analysis of the relatinship between displacement and dimensions of faults. *J. Struct. Geol.*, **10**, 239-247.

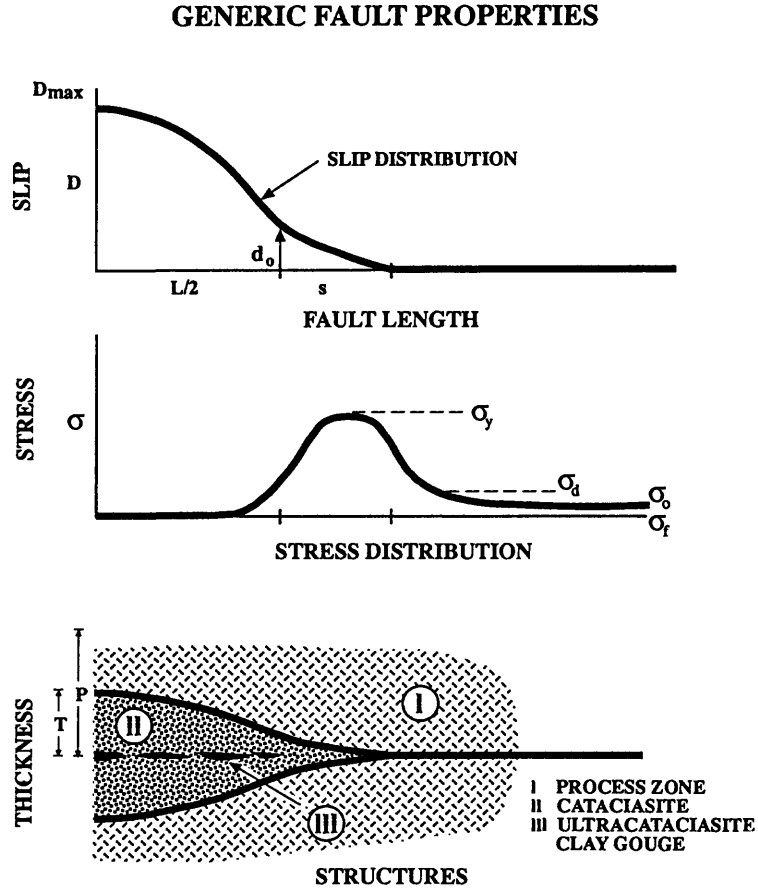



Fig. 1. A generic fault model, as described in text.

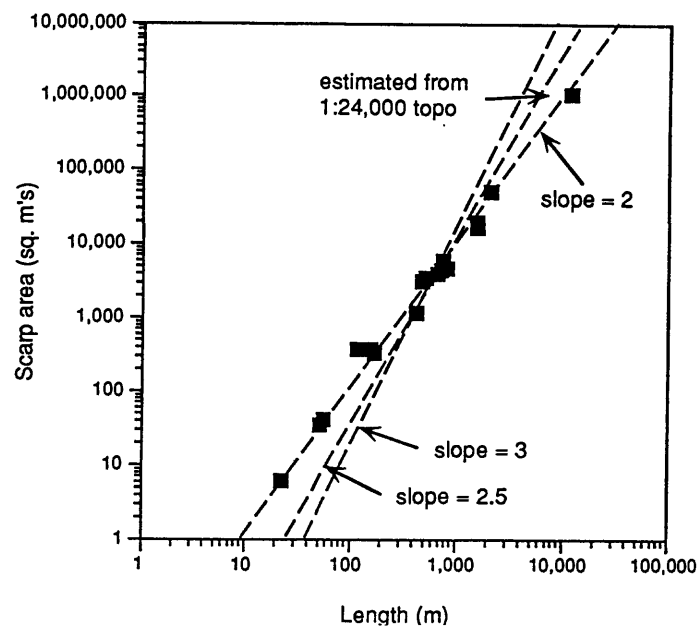



Fig. 2. Scarp area vs. fault length for a suite of normal faults, Volcanic Tablelands, Eastern California. The line of slope 2 indicates linear scaling between displacement and length.

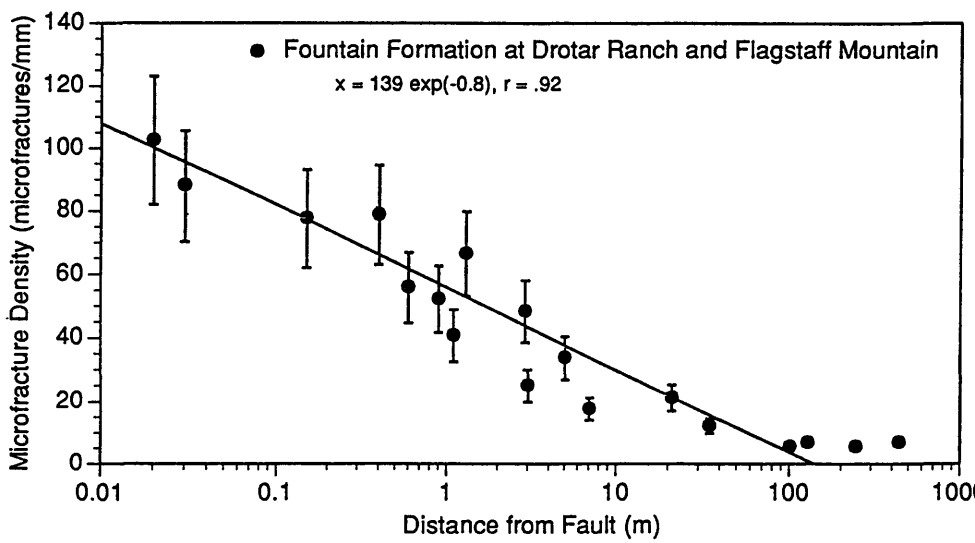



Fig. 3. Microcrack density as a function of distance from faults: two faults cutting the same rock type, Colorado (from Anders and Wilschko, 1993).

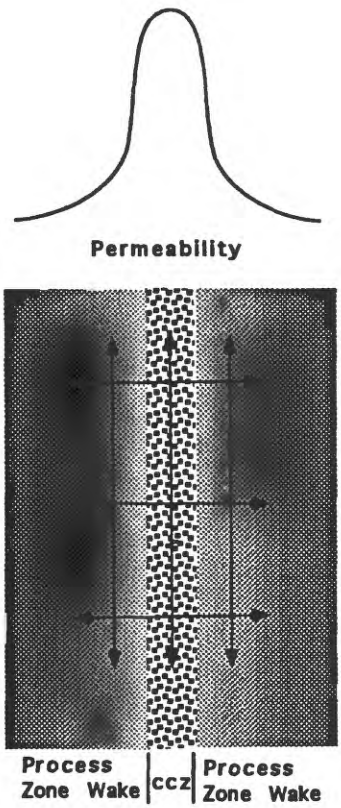



Fig. 4a Open cataclasite zone (ccz), open process zone. Fault is permeable conduit both parallel and transverse to strike. Case where protolith is clay free and fault zone is unmineralized.

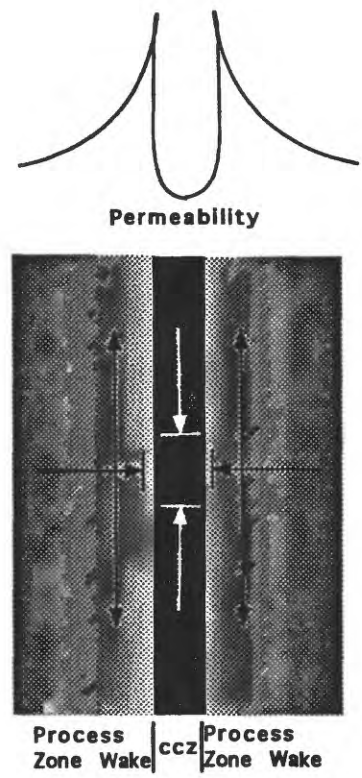



Fig. 4b. Outer conduit. Cataclasite zone is clay rich. Process zone is unmineralized.Fault is permeable along strike \in process zone and impermeable transverse to strike.

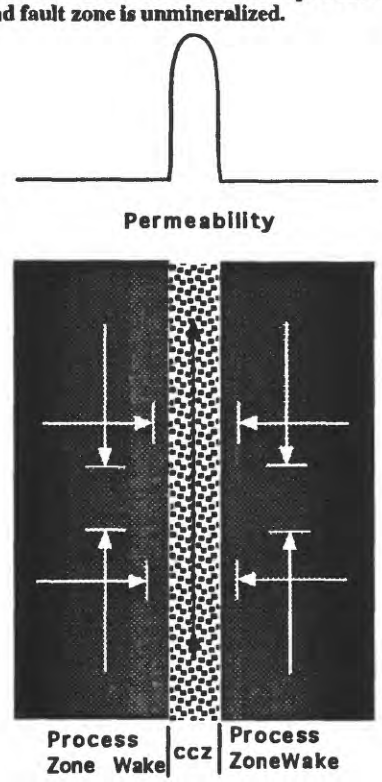



Fig. 4c. Inner conduit. Active fault, process zone mineralized, cataclasite zone refractured. Transport only within ccz.

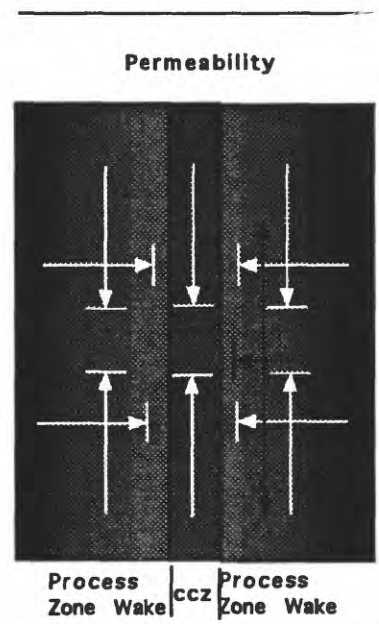



Fig. 4d. Inactive fault: both cataclasite zone and process zone mineralized. Impermeable both parallel and transverse to strike.

# The role of water in the evolution of large crustal faults such as the San Andreas in California

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

J. Byerlee

Studies of crustal permeability by Brace (1980), studies of isotope data from several intrusive bodies by Taylor (1990) and similar studies of regional metamorphic systems by Walther (1990), indicate that in most regions the crust is saturated with meteoric water at hydrostatic pressure. When a fault is formed water will flow into the highly porous and permeable fault zone until the pore pressure in the fault is equal to the hydrostatic pressure in the country rock. In a strike-slip regime the maximum principal stress $\sigma_1$, must be greater than the lithostatic stress and in California $\sigma_1$ is almost normal to the San Andreas fault. Therefore, at a depth of 15 km, the base of the seismogenic zone, the effective normal stress $\bar{\sigma}_n$ must, at least initially, be greater than 200 MPa.

At this high pressure the initially highly porous material in the fault zone will become crushed and compacted. Because of the large specific surface area of the finely crushed gouge and the high temperature that exists at this depth the ultra fine particles in the gouge will react rapidly with the surrounding material. The solubility of the particles is given by the Ostwald-Freundlich equation $Sr = Si$ exp (2EV/RTr) where $Sr$ is the solubility of particles of radius $r$, $Si$ is the solubility of a flat surface or a particle of infinite radius, $E$ is the interfacial energy, $V$ is the molar volume, $R$ is the gas constant, and $T$ is the absolute temperature (Iler, 1978).

The geometry of a fault zone varies greatly both in the vertical direction and along the fault, but in general there are one or more zones of fault gouge that give way to a breccia zone that in turn is separated from the country rock by a fracture zone (Wallace and Morris, 1986). During compaction of the fault zone, water saturated with silica in equilibrium with the fine particles will be forced back into the country rock through the breccia and fracture zone where it will be oversaturated with respect to the more-or-less planar surfaces in this new environment. Deposition of silicates will reduce the width of the passageways through which the water moves, resulting in a decrease in the permeability of the outer margins of the fault zone.

Because $\bar{\sigma}_n$, at least initially, and the temperature increase with depth, the time dependent compaction of the dilated fault zone will occur most rapidly at the deepest part of the fracture zone, and the water in the pore spaces will be forced upward through the fault zone. The water will become oversaturated with silica as it moves into the low temperature environment, and this will facilitate the deposition of silicates in the walls of the fault as well as in the fault zone itself, if it encounters heterogeneities in either the structure or the mineralogical composition of the fault.

One of the recent advances in material sciences is the discovery that a liquid has finite strength if the separation of the walls between which the liquid is confined is less than 10 times the diameter of the liquid molecules (Israelachvili *et al.*, 1988). Because of this, flow of water through a material will stop altogether if the pore pressure gradient falls below a threshold value (Byerlee, 1990). Therefore, fluid flow will decrease if the width of

the channelways through which the water moves is reduced because of silicate deposition to less than 10 times the diameter of a water molecule. The high pore pressure in the fault zone can be maintained indefinitely if the pressure gradient falls below a threshold value. In other words, seals form between the fault zone and the country rock and within the fault zone itself without the necessity of completely clogging the system with silicates. When flow stops the narrow passageways begin to heal, leaving behind an array of fluid-filled inclusions. In this way the seals become almost impermeable and water can migrate through them only by molecular diffusion. The result of this process will be the formation of a three dimensional network of seal-bounded fluid compartments having no hydraulic communication with each other or with the external hydraulic regime in the country rock as shown schematically in Figure 1.

The theoretical models (Byerlee, 1990, Byerlee, 1992, Rice, 1992) show that in an actively deforming fault zone the pore pressure at any point is less than the local minimum principal stress $\sigma_3$ provided that the pore fluid penetrates the material to lower the effective stress. When an impermeable seal is formed, this condition is not satisfied and the now lithified material will stop shearing; therefore, $\sigma_3$ in the seal will remain at the stress state established during the last shearing episode. If compaction of the material in the compartment continues after the seal is formed, a pressure will be reached where the now-brittle seal will undergo hydrofracture, fluid will escape into the surrounding rock, and the pore pressure in that compartment will decrease. Because of the change in $\sigma_n$ acting on the drained section of the fault it will compact further until the fractures once more seal by deposition of silicates. Because the the pore pressure is now lower in the compartment, $\bar{\sigma}_n$ across the fault in this section will increase and the fault will become locked.

To accommodate the movement of the crustal plates, the shear stress on this locked section of the fault will increase until it fails in shear. However, the first place to fail will be the relatively rigid, lithified seal between this low-pressure, low-porosity compartment and adjacent high-pressure, high-porosity compartment. This is shown schematically in Figure 2, a horizontal cross section of two of the compartments shown in the vertical section in Figure 1. Here $A$ is the area of the compartment, $W$ is its width, $\eta$ is its porosity, and $P$ is the pressure minus the hydrostatic gradient of the fluid in the compartment; the subscripts refer to compartments 1 and 2, respectively.

When the seal between the two compartments is ruptured, water will flow from the high-pressure to the low-pressure compartment. If we neglect the energy lost through frictional resistance of the fluid moving through pore spaces, then the energy lost in moving water of mass $m$ from the high-pressure, high-porosity compartment is gained when it moves into the low-pressure, low-porosity compartment. This is given by the equation

$$\int_0^m \int_{P_1}^{P_f} V_1 \, dpdm + \int_m^0 \int_{P_2}^{P_f} V_2 \, dpdm = 0, \tag{1}$$

where $V = \eta AW$ is the pore volume in the compartments.

Assuming that the compressibility of the water remains nearly constant, then equation 1 leads to the expression

$$P_f = \left( \frac{V_1 P_1 + V_2 P_2}{V_1 + V_2} \right). \tag{2}$$

The average initial pressure $\bar{P}_{in}$ in the two compartments is given by the equation:

$$\bar{P}_{in} = \left( \frac{P_1 A_1 + P_2 A_2}{A_1 + A_2} \right). \tag{3}$$

If $P_f > \bar{P}_{in}$ the average effective normal stress across the two compartments will be lower in the final state than it was initially and the shear strength decrease, $\Delta\tau$, from the initial to the final state is given by the equation

$$\Delta\tau = \mu \left[ \left( \frac{P_1 V_1 + P_2 V_2}{V_1 + V_2} \right) - \left( \frac{P_1 A_1 + P_2 A_2}{A_1 + A_2} \right) \right], \tag{4}$$

where $\mu$ is the coefficient of friction. If $P_2 > P_1$ then $\Delta\tau$ is positive if $\eta_2 W_2 > \eta_1 W_1$.

If the pore fluid is contained in seal-bounded fluid compartments, and it is in hydraulic communication within each compartment, then the pressure may be lithostatic at the top of a compartment, but it must decrease along the hydraulic gradient to the bottom of the compartment. In an adjacent compartment, the pressure may be hydrostatic everywhere as shown schematically in Figure 3. The pressure will not be evenly distributed in the fault zone, but the fault zone will be weakened to such an extent that the heat flow and stress direction restrictions on a fault such as the San Andreas will be satisfied if the pressure is hydrostatic in the country rock and the average pressure at the center of the fault is ≈85% of lithostatic at all depths (Byerlee, 1992).

When the seal between the two compartments is broken either by tectonic forces or transient forces generated during the passage of a seismic wave from a distant earthquake, water will diffuse from the high pressure to the low pressure compartment. If the permeability $K_2 >> K_1$ the pore pressure will be uniform throughout compartment 2. The time for the average pressure in compartment 1 to reach about 90% of its equilibrium value is a function of $\eta_1 L_1^2 \beta\nu/K_1$ and the ratio of $V_2/V_1$ where $\beta$ is the compressibility and $\nu$ is the viscosity of water at the prevailing temperature and pressure (Crank, 1986). The time dependent decrease in the shear strength of the two compartments is shown in Figure 4 for two values of the ratio of $V_2/V_1$. An earthquake will be triggered when the average shear strength of the two compartments is decreased to a value less than the tectonic shear stress on this section of the fault.

When the fault suddenly slips, the geometrical irregularities on the fault will be fractured and the gouge, breccia zone and fracture zone will increase in width and the process described above will be repeated. Thus, in this model the process is an episodic one with water moving in and out of the fault zone like the seismic pumping mechanism first described by Sibson *et al.* (1975). In our model high pore pressure is generated in the fault zone by time-dependent compaction in the seal-bounded compartments. Thus, the high fluid pressures are self-generating.

REFERENCES

Brace, W.F., 1980, Permeability of crystalline and argillaceous rocks: Status and problem: International Journal of Rock Mechanics, v. 17, p. 876–893.

Byerlee, J., 1990, Friction, overpressure and fault normal compresson: Geophysical Research Letters, v. 17, p. 2109–2112.

Byerlee, J., 1992, The change in orientation of subsidiary shears near faults containing high pore fluid pressure: Tectonophysics, v. 211, p. 295–303.

Crank, J., 1986, The mathematics of diffusion: Oxford, Clarendon Press, 414 p.

Iler, R.K., 1978, The chemistry of silica; solubility, polymerization, colloid and surface properties, and biochemistry: New York, John Wiley and Sons, 787 p.

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

Rice, J.R., 1992, Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault, *in* Evans, B., and Wong, T.F., eds., Earthquake mechanics and transport properties of rocks: London, Academic Press, p. 475–503.

Sibson, R.H., Moore, J.M. and Rankin, A.H., 1975, Seismic pumping - A hydrothermal fluid transport mechanism: Geological Society of London Journal v. 131, p. 653–659.

Taylor, H.P., 1990, Oxygen and hydrogen isotope constraints of the deep circulation of surface waters into zones of hydrothermal metamorphism and melting, in Bredehoft, J.D., and Morton, D.L., eds., *The role of fluids in crustal processes*: Washington, D.C., National Academy Press, p. 72–95.

Wallace, R.E., and Morris, H.T., 1986, Characteristics of faults and shear zones in deep mines: Pure and Applied Geophysics, v. 124, p. 107–125.

Walther, J.F., 1990, Fluid dynamics during progressive regional metamorphism, *in* Bredehoeft, J.D., and and Morton, D.L., eds., The role of fluids in crustal processes: Washington, DC, National Academy Press, p. 64–71.

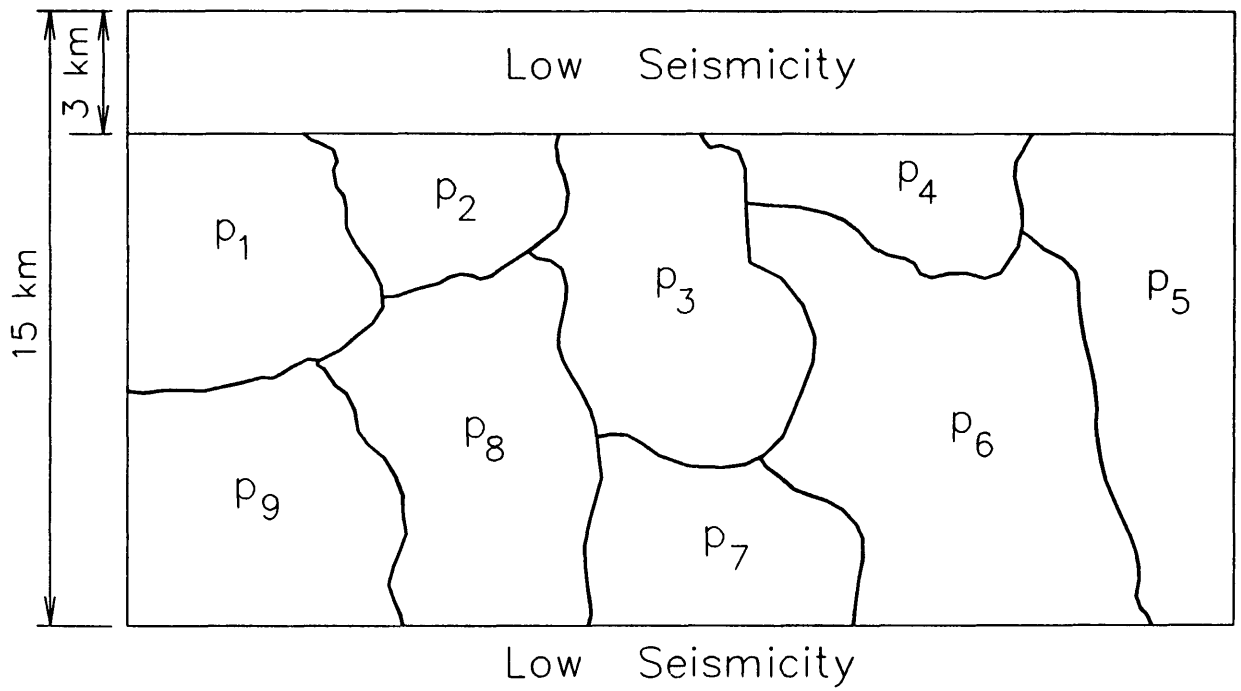



Figure 1. Schematic diagram of vertical section along a large crustal fault such as the San Andreas. Pressure $P_i$ in each compartment will be different because compartments are separated by impermeable seals.

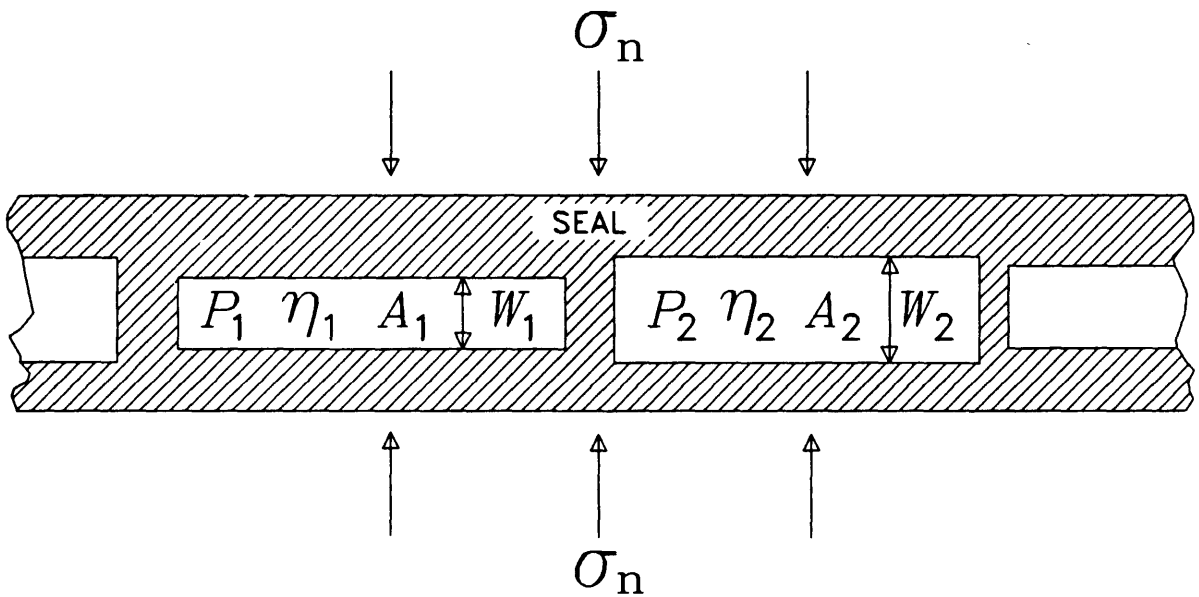



Figure 2. Horizontal cross section through two of the compartments illustrated in Figure 1. $P$ is pore pressure; $\eta$ is porosity; $A$ is area; and $W$ is width of compartments.

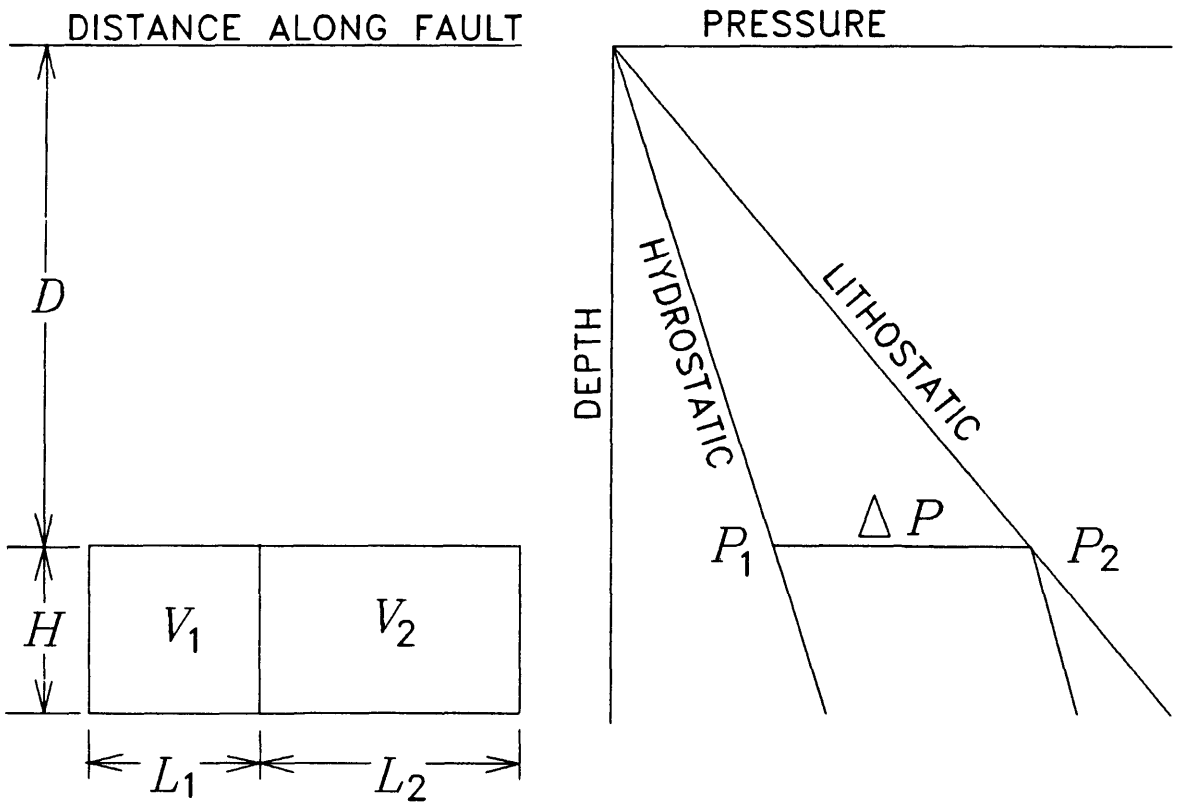



Figure 3. Schematic diagram of compartments 1 and 2. $P_1$ is hydrostatic everywhere, and $P_2$ is lithostatic at the top and increases along hydrostatic pressure gradient to bottom.

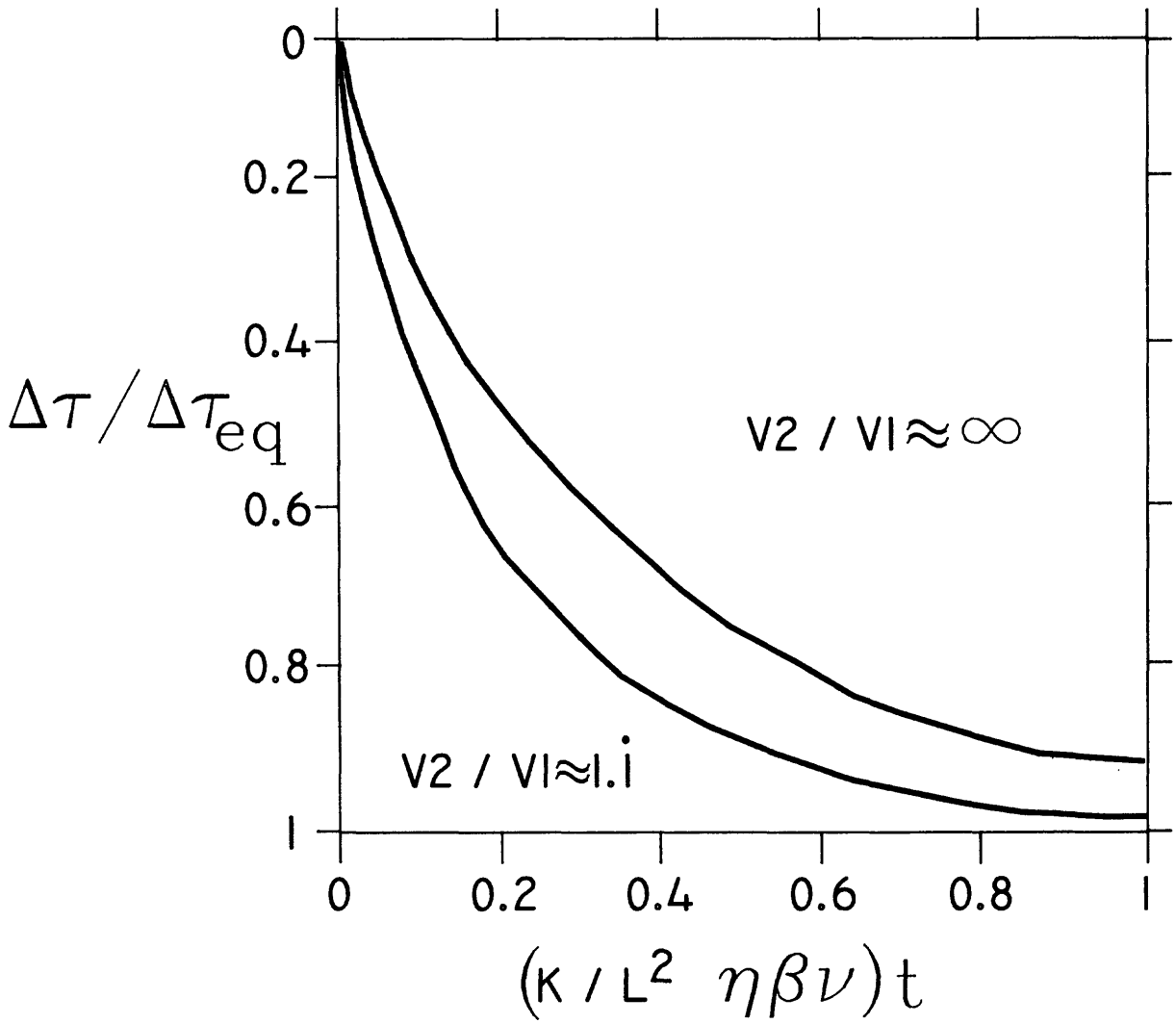



Figure 4. Shear strength decrease $\Delta\tau$ of the two compartments normalized by the equilibrium shear strength decrease $\Delta\tau_{eq}$, as a function of time normalized by $\eta L^2 \beta\nu / K$.

# MAGNETIC AND ELECTRIC FIELDS ASSOCIATED WITH CHANGES IN HIGH PORE PRESSURE IN FAULT ZONES - APPLICATION TO THE LOMA PRIETA ULF EMISSIONS.

By M.A. Fenoglio, M. J. S. Johnston and J.D. Byerlee

## ABSTRACT

We determined the electric and magnetic fields generated during failure of faults containing sealed compartments with pore pressures ranging from hydrostatic to lithostatic levels. Exhumed fault studies and strain measurement data limit the possible size of these compartments to less than 1 km in extent. Rupture of seals between compartments produces rapid pore pressure changes and fluid flow and may create fractures that propagate away from the high pressure compartment, along the fault face. Non-uniform fluid flow results from pressure decrease in the fracture caused by crack generated dilatancy, partial blockage by silicate deposition, and clearing as pressure increases. The direct consequences of this unsteady fluid flow are associated transient magnetic signals caused by electrokinetic, piezomagnetic, and magnetohydrodynamic effects. Models of these processes for fault geometries with 1 km high pressure compartments show that electrokinetic effects are several orders of magnitude larger than the other mechanisms. The electrokinetic signals produced by this unsteady flow are comparable in magnitude and frequency to the magnetic signals observed prior to the $M_L$ 7.1 Loma Prieta earthquake of 18 October 1989 provided fracture lengths are less than about 200 m.

## INTRODUCTION

Byerlee (1990, 1992) and Rice (1992) recently suggested a resolution of the "San Andreas Heat Flow Paradox" by proposing a model in which fluid pressure gradients within the fault zone can produce a fault with low strength while avoiding hydrofracture in the surrounding rock due to excessive fluid pressure. Byerlee (1993) extended this idea by suggesting that silicate deposition within the fault zone and between the fault zone and the country rock results in the formation of sealed compartments of various sizes and porosities. Compaction increases the fluid pressure to levels that are greater than hydrostatic but not uniform from compartment to compartment along the fault.

The formation, growth and failure of these high pressure compartments has interesting consequences, particularly for fault mechanics and earthquake prediction research. Changing pore pressure from perhaps lithostatic to hydrostatic in these inclusions should, depending on their size and location, generate readily detectable strain changes, microseismicity, and electric and magnetic field transients prior to fault failure. Observations of high resolution strain before moderate earthquakes (Johnston et al., 1987), microearthquake patterns (Nadeau et al., 1993), and indications from exhumed faults, suggest characteristic compartment sizes could be as large as 100 m to 1000 m in extent. Aseismic failure of these inclusions could provide the trigger mechanism for larger-scale sonic fault rupture (conventional earthquakes).

Our principle interest here concerns magnetic and electric fields that might result from mechanisms initiated by rapid pore pressure changes and fluid flow following rupture or partial rupture of these compartments. These mechanisms include electrokinetic effects (Mizutani, 1976; Fitterman, 1978,1979; Ishido and Mizutani, 1981; Miyakoshi, 1986; Dobrovolsky et al., 1989), piezomagnetic effects (Stacey, 1964; Stacey and Johnston, 1972; Davis et al., 1979; Sasai, 1980; Davis et al., 1984; Johnston, 1989), and possible contributions from magnetohydrodynamic effects (Draganov et al., 1991) and stress-resistivity changes (Mazella and Morrison 1974; Fitterman and Madden, 1977; Madden, 1979; Morrison et al., 1979; Qian et al., 1983; Sheng and Chen, 1988; Park and Fitterman, 1990). In this paper we investigate comparative electrokinetic, piezomagnetic, and magnetohydrodynamic models, and realistic extensions of these models, expected from the simple "compartment failure" hypothesis. In particular, we suggest that magnetic fields resulting from irregular fluid flow should be expected

given the dramatic physical changes during compartment failure. Further, these effects at the hypocenter of the 1989 $M_L$ 7.1 Loma Prieta earthquake may explain the increased ULF noise reported by Fraser-Smith et al (1990) near the epicenter before and after the earthquake.

## HIGH PORE PRESSURE FAULT MODEL

In Byerlee's (1993) model of high fault zone pressure, impermeable seals hydrologically isolate the fault zone from the surrounding country rock. A three-dimensional network of impermeable seals within the fault zone confines fluids at suprahydrostatic to lithostatic pressures in compartments of various sizes and shapes. The size of these compartments is poorly constrained. Observations of fluid inclusions with high apparent pore pressure in exhumed faults provide some indication of compartment size. Kerrich et al. (1977) observes compartment sizes of up to 1 km. For currently active compartments, size may be inferred indirectly through near-surface strainmeter data. Strainmeters in 200-m deep boreholes near the San Andreas fault in California operate at sensitivities of about $10^{-9}$. Since strain changes above this measurement limit are not readily apparent in these data before large earthquakes (Johnston et al, 1987), compartments at seismogenic depth could be as large as several hundred meters if rapid pressure changes occur at the kilobar level. For these reasons, we initially assume compartments have dimensions less than 1 km.

Equally important to this model is the pore pressure distribution in the fault zone. We may infer a probable pore pressure system from the heat flow measurements along the San Andreas fault. Byerlee (1992) provides an expression which relates the apparent coefficient of friction with the ratio of the pore pressure in the fault zone to the magnitude of the vertical stress. For a coefficient of friction, $\mu < 0.1$, implied by the absence of a heat flow anomaly over the fault, the average pore pressure in the fault zone must be 85% of lithostatic. If the fault zone consists exclusively of lithostatic and hydrostatic pore pressure compartments, 75% of the fault zone volume must be lithostatic pore pressure compartments. In other words, lithostatic compartments must contain three times the volume of hydrostatic compartments.

During differential plate motion, shear stress builds in the weak, fluid-filled fault zone until the strong impermeable seals between high- and low-pressure compartments rupture. Rupture of seals in response to tectonic loading would initiate shear fractures that propagate into the low-pressure compartment, enabling the rapid movement of fluid from the high-pressure compartment. Figure 1 illustrates the geometry of this process.

For this model, we consider a relatively thin shear fracture, in which fluid is constrained to flow primarily along fault strike, in the direction of fracture propagation. The expression for one-dimensional transient flow of a compressible fluid through a porous medium is (Brace et al, 1968)

$$\frac{\partial^2 P}{\partial x^2} = \frac{\eta\beta}{k}\left[\frac{\beta_{eff} - \beta_s}{\beta} + \varphi\left(1 - \frac{\beta_s}{\beta}\right)\right]\frac{\partial P}{\partial t} \tag{1}$$

where P is pressure in the fracture, $\eta$ is dynamic viscosity, k is fracture permeability, $\varphi$ is the porosity, $\beta$ is the fluid compressibility, $\beta_S$ is the compressibility of the solid matrix, and $\beta_{eff}$ is the effective compressibility of the rock (as measured in a jacketed sample). For most rocks, $\beta$ is much larger than either $\beta_S$ or $\beta_{eff}$ (Brace, 1968). Therefore, the flow expression becomes

$$\frac{\partial^2 P}{\partial x^2} = \frac{\eta\beta\varphi}{k}\frac{\partial P}{\partial t} \tag{2}$$

The shear fracture propagates into the low-pressure compartment. At any time in its evolution, the fracture is dilatant. Exhumed fault studies support this argument (Sibson et al, 1988; Chester and Logan, 1986). Thus, the initial condition of the system P(x,0) is small since

only the fracture tip is present at the time of fracture initiation. Further, since the volume of the shear fracture is negligible compared to the volume of the high-pressure compartment, the pore pressure at the compartment boundary (where rupture first occurs) remains at the high-pressure levels for the duration of the process. In our extension of the Byerlee model, we assume the high-pressure compartment has lithostatic pore pressure levels. To simplify our mathematical analysis, we use the method of images, which essentially doubles the width of the low-pressure compartment, with a high-pressure compartment on either side. This is equivalent to reflection about the y-axis in Figure 1. Thus, the boundary condition is $P(\pm\ell, t) = P_L$.

Here we assume that $P \rightarrow 0$ as $x \rightarrow \infty$. This provides an approximation of the real behavior of the system. Applying Laplace transforms yields a pore pressure distribution in the shear fracture of

$$P(x,t) = P_L - P_L \operatorname{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{3}$$

where $D = k/\eta\beta\varphi$. The pore pressure distribution is illustrated in Figure 2. In a more realistic description of the shear fracture evolution, $P \rightarrow 0$ as $x \rightarrow \ell$, and the average pore pressure in the fracture is hydrostatic. Equation (3) above, because of the differences in formulation, is only an approximation of the more realistic description.

This pore pressure distribution constantly evolves as the fracture propagates further into the low pressure compartment. Since the pore pressure distribution in the fracture varies from lithostatic to near-zero pressures at fracture tip, deposition of material will occur as the solubility, which is strongly pressure dependent, changes (Fournier, 1985; Herrington and Wilkinson, 1993; Fournier and Potter, 1982). At a depth of 17 km, the silica solubility across the fracture drops from 1% weight percentage near the lithostatic high-pressure compartment to less than 0.5% weight percentage at the fracture tip (see Figure 3). Consequently, silica deposition near the fracture tip may decrease permeability, reducing and perhaps stopping further fluid flow and fracture propagation. The precise amount of silica deposition necessary is the result of complications in the silica deposition rate, the geometry of the shear fracture, the amount of gouge material present in the fracture, and the initial quantity of silica present in the fluid. Should sufficient deposition occur to form a temporary seal, pressure builds in the fracture until rupture of the new seal occurs and the shear fracture propagation continues. This stop-and-start fracture propagation continues until the shear fracture is stopped or it reaches another high-pressure compartment. In the latter case, fluid will rush in from the newly breached high-pressure compartment and the pore pressure distribution in the fracture becomes quite complex. The shear fracture may continue to propagate along the fault through the high pressure compartment, producing seismic radiation and moving at Rayleigh wave velocity (Andrews, 1976). Should this fracture reach another impermeable seal between a high-pressure and low-pressure compartment, the seal ruptures and the entire process begins anew. Each time a seal ruptures or a shear fracture stops and starts, fluid motion occurs, possibly generating observable magnetic and electric field changes.

## ELECTROKINETIC EFFECTS

Electrokinetic effects are the electrical currents (and magnetic fields) generated by fluid flow through the crust in the presence of an electric double layer at the solid-liquid interfaces (Overbeek, 1953). This double layer consists of ions anchored to the solid phase, with an equivalent amount of ionic charge of opposite sign distributed in the liquid phase near the interface. For a shear fracture into which fluid flows, this electric double layer exists at the fracture wall.

When the fluid in such a system moves, the charges in the fluid are transported along the fracture wall in the direction of fluid motion; that is, electric current exists. The current density

and fluid flow are coupled processes described by the following equations (Nourbehecht, 1963; Fitterman, 1979)

$$\mathbf{j} = -\sigma \nabla \Psi - \frac{\varepsilon \zeta}{\eta} \nabla P \tag{4}$$

and

$$\mathbf{v} = -\frac{\varepsilon \zeta}{\eta} \nabla \Psi - \frac{k}{\eta} \nabla P \tag{5}$$

where $\mathbf{j}$ is the current density, $\mathbf{v}$ is the fluid velocity, $\Psi$ is the streaming potential, $\varepsilon$ is the dielectric constant, $\zeta$ is the zeta-potential, and $\sigma$ is the fluid conductivity. The zeta-potential is the measure of the initial potential (which exists because of a non-uniform distribution of the ions in the fluid phase) in the electric double layer.

These expressions imply that an imposed electric field can initiate fluid motion. Probstein and Hicks (1993) demonstrated this effect by using a direct current electric field to induce motion of fluid in the earth. The converse of the process is fluid motion generating an electric current. Clearly, an absence of fluid motion indicates a lack of electric current; that is, the system is in equilibrium. A potential still exists in the form of a non-uniform distribution of ions in the fluid, but there is no movement of these ions.

In modeling our fracture, then, we have large pressure gradients driving fluid flow, as governed by the expression for pore pressure within the shear fracture (3). Because of the large pressure gradients involved, the second term in equation (4) is dominant, so

$$\mathbf{j} \approx -\frac{\varepsilon \zeta}{\eta} \nabla P$$

This simplifies our electrokinetic construct. We simply say the fluid within the fracture flows at approximately the same rate as the fracture propagates. Thus, when the fracture ceases propagation (and the fracture is entirely fluid filled; that is, soon after another high-pressure compartment is reached), the fluid flow will quickly cease and the electrokinetic effects end. The systems has returned to equilibrium. By definition, this equilibrium has a fluid flow of zero as a necessary condition. Proceeding from these arguments, our current density is

$$\mathbf{j} = \hat{a}_x \cdot \frac{P_L \varepsilon \zeta}{\eta \sqrt{Dt}} e^{-x^2/4Dt} \tag{6}$$

This current density in the shear fracture generates a magnetic field at the surface. We apply the Biot-Savart law to determine the magnetic field.

$$\mathbf{B} = \frac{\mu_o}{4\pi} \int_A \frac{\mathbf{j} \times \mathbf{r}}{r^2} dA = \frac{\mu_o j_o}{4\pi \sqrt{t}} \int_0^t \int_d^{d+h} \left( -\frac{z - z'}{r^3} \hat{\mathbf{a}}_y + \frac{y}{r^3} \hat{\mathbf{a}}_z \right) \exp\left[ \frac{-(x - x')^2}{4Dt} \right] dz' \, dx' \tag{7}$$

where $r = \sqrt{(x - x')^2 + y^2 + (z - z')^2}$, $j_o = P_L \varepsilon \zeta / \eta \sqrt{D}$, and $\mu_0$ is the magnetic permeability of free space. The magnetic field cannot be determined analytically from this integral expression. We can integrate with respect to the vertical variable (z') to reduce the problem to a single integral. Using integration by parts allows numerical integration that reveals the horizontal

magnetic field ($B_y$) is several orders of magnitude larger than the vertical magnetic field ($B_z$). Therefore, we will limit our analysis to the horizontal magnetic field at the surface.

For a current sheet 200 m in length and 1000 m in height, the surface magnetic field can be calculated for different depths and times using the physical constants summarized in Table 1. Figures 4 and 5 show the magnetic field profiles along fault strike for varying depths and times, respectively. As the depth of the source increases, the peak magnetic value decreases and the signal becomes more and more spatially distributed. As time increases, the magnetic field decreases rather uniformly along fault strike. The curve labeled t = 1 sec in Figure 5 corresponds to the curve labeled z = 17 km in the Figure 4.

For the modestly sized source considered, the magnetic field at the surface may be as high as ±2 nT even for a source buried at 17 km. Clearly, electrokinetic effects resulting from rupture of impermeable layers between fault compartments of differing pore pressure may be quite large.

## MAGNETOHYDRODYNAMIC EFFECTS

The motion of an electrically conducting fluid in the presence of a magnetic field perturbs that field. We begin with Maxwell's equations

$$\nabla \times \mathbf{B} = \mu_o \mathbf{j} \tag{8a}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \tag{8b}$$

and Ohm's law

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{9}$$

where **B** is the magnetic field (T), **E** if the electric field (V/m), and **j** is the electric current density ($A/m^2$). If we substitute Ohm's law (9) into the first of Maxwell's equations (8a) and take the curl of both sides,

$$-\nabla^2 \mathbf{B} = \mu_o \sigma(\nabla \times \mathbf{E}) + \mu_o \sigma \nabla \times (\mathbf{v} \times \mathbf{B})$$

Using the second of Maxwell's equations yields the magnetohydrodynamics (MHD) equation,

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \lambda \nabla^2 \mathbf{B} \tag{10}$$

where $\lambda = 1/\mu_o\sigma$ is the magnetic diffusivity ($m^2/s$). This expression illustrates the two terms which define the character of the magnetic field. The first term on the right is the convection of the magnetic field caused by the resistance to change of the flux linked by a closed, highly conductive loop. The second term is diffusion of the magnetic field caused by ohmic dissipation (Shercliff, 1962).

This equation is analogous to the dynamic expression of viscous, non-conducting fluid. In such a system, a hydrodynamic Reynolds number defines the relative importance of the convection and diffusion terms. We may employ a magnetic Reynolds number to make the same comparison for the MHD expression. Dimensional analysis of the convective and diffusion terms gives us

$$\frac{|\nabla \times (\mathbf{v} \times \mathbf{B})|}{|\lambda \nabla^2 \mathbf{B}|} \approx \frac{|\mathbf{v} \times \mathbf{B}|}{\lambda |\nabla \times \mathbf{B}|} \approx \frac{vB}{\lambda B / \ell} = \frac{v\ell}{\lambda} = \mu_o \sigma v \ell$$

where $\ell$ is a characteristic length of the source. This ratio is the magnetic Reynolds number $R_m$ that defines the amplitude of an induced field produced by motion of an electrically conductive fluid in an applied field. For our problem, the applied field is simply the Earth's magnetic field and the other characteristics are defined by the fluid action in the fault zone.

The fluid velocity in the shear fracture can be calculated approximately from the pore pressure distribution in equation (3), shown in Figure 2. The expression (3) is a complimentary error function which asymptotically approaches zero as the horizontal variable approaches infinity. The region which covers 95% of the pore pressure distribution is $0 < x/2\sqrt{Dt} < 1.36$ (see Figure 2). This area represents an approximation of the finite-length fracture. Under this assumption, we can determine an approximate form for fluid velocity in the shear fracture.

$$\frac{x}{2\sqrt{Dt}} = 1.36$$

The fluid velocity, v, is

$$v = \frac{dx}{dt} = \frac{3.70D}{x}$$

There is a singularity in fluid velocity at the point of rupture between the high- and low-pressure compartments. However, since the fluid velocity cannot exceed the crack propagation rate and since the maximum crack propagation rate is the Rayleigh wave velocity (Andrews, 1976), the peak flow velocity is the Rayleigh wave velocity. Eberhart-Phillips et al (1990) and Barker et al (1990) investigated seismic wave velocities associated with the Loma Prieta earthquake. At 17 km depth, the P-wave velocity was about 6.5 km/sec. For a Poisson solid, the Rayleigh wave velocity at the same depth is about 3.5 km/sec. This velocity corresponds to a distance along the fracture of $x = 1.1 \times 10^{-3}$ D. Hence, to eliminate the velocity singularity, we assume the fluid moves at Rayleigh wave velocity for $0 < x < 1.1 \times 10^{-3}$ D. For the values in Table 2, this corresponds to a distance of about 7 cm.

The MHD effect from the fluid motion within the fracture is

$$B(x) = R_m B_o = \mu_o \sigma v \ell B_o = \mu_o \sigma \ell B_o \begin{cases} v_R, & x < 1.1 \times 10^{-3} D \\ \dfrac{3.7D}{x}, & x \geq 1.1 \times 10^{-3} D \end{cases} \qquad (11)$$

where $\mu_o$ is the magnetic permeability of free space and $B_o$ is the Earth's magnetic field (about $5 \times 10^{-5}$ T). For reasonable values of the relevant physical parameters (see Table 2), the maximum magnetic field from the MHD effect is approximately 4900 nT at the source. If the source in the fracture acts as a magnetic dipole, the surface magnetic field from the MHD effect is negligible (given the $1/r^3$ falloff in signal strength).

This result differs from that obtained by Draganov et al (1991), who proposed an MHD model for the region surrounding the Loma Prieta earthquake in which a conductive sheet of water at 4 km depth moves with a fluid velocity of 4 cm/s. With this model he calculated a surface magnetic field of approximately 0.1 nT. The difficulties with Draganov's model concern the permeability of $10^{-10}$ m$^2$ proposed for this layer (i.e., equivalent of sand), the pressure gradient required to drive fluids at this velocity, the non-realistic nature of sheet fluid flow with its implications for strain, seismicity, and surface hydrology, and the need to make this flow

oscillatory with a wave number, $k_x$, of $10^{-4}$, if this model is to approximate the electromagnetic signals observed by Fraser-Smith et al (1990) for this earthquake. Even supposing a fracture system exists in this region, the permeability is unlikely to exceed $10^{-11}$ $m^2$ (Brace, 1980) under the most favorable of circumstances. With this permeability, the pressure gradient needed is $4 \times 10^6$ Pa/m. For this case, then, the pore pressure would be

$$P = \frac{1}{k_x}\frac{dP}{dx} = 4 \times 10^{10}\ Pa$$

This pore pressure is well above lithostatic (about $10^8$ Pa at 4 km depth) and is extremely unlikely to exist. We therefore conclude that magnetohydrodynamic effects at depth appear not to contribute to anomalous surface magnetic fields.

## PIEZOMAGNETIC EFFECTS

The piezomagnetic effect results from a change in magnetization of ferromagnetic-bearing rock in response to applied stress. The general formula for the piezomagnetic effects for isotropic magnetic materials is given by (modified from Sasai, 1991)

$$\Delta M_i = \left(-\tfrac{1}{2}\tau_{kk}\delta_{ij} + \tfrac{3}{2}\tau_{ij}\right)\beta M_j \tag{12}$$

where $\Delta M$ is the change in magnetization, $\beta$ is the stress sensitivity, $\tau$ is the stress tensor, and $\delta_{ij}$ is the Kronecker delta function. If the material obeys Hooke's law,

$$\tau_{mn} = \lambda\delta_{mn}\nabla\cdot\mathbf{u} + \mu\left(\frac{\partial u_m}{\partial x_n} + \frac{\partial u_n}{\partial x_m}\right) \tag{13}$$

where $\lambda$ and $\mu$ are Lamè's constants and $\mathbf{u}$ is the displacement vector, we can derive an equation which relates displacement components to piezomagnetism by substitution of Hooke's law into our isotropic piezomagnetic constitutive relation.

$$\Delta M_{kl} = \left[-\delta_{kl}\nabla\cdot\mathbf{u} + \tfrac{3}{2}\left(\frac{\partial u_m}{\partial x_n} + \frac{\partial u_n}{\partial x_m}\right)\right]\mu\beta M_k \tag{14}$$

Here, $\Delta M_{kl}$ is the $l^{th}$ component of the piezomagnetic effect produced by the $k^{th}$ component of the initial value (Sasai, 1991).

As the shear fracture propagates to another high-pressure compartment, the high pore pressures which act along the walls of the fracture induce a piezomagnetic effect. We model this effect by treating the fracture as a dyke subject to stress equal to the pore pressure of the high-pressure compartment. Using the difference equation (14), we can numerically integrate to determine the magnetic field at the surface which results from piezomagnetic effects (Sasai, 1991).

Table 3 summarizes the physical constants and fracture geometry assumed in the numerical integration. The magnetic field at the surface is everywhere less than $10^{-2}$ nT, and is strongest directly over the source. Piezomagnetic effects, while significant compared to any MHD contribution, are still two orders of magnitude less than electrokinetic effects.

## DISCUSSION OF THE LOMA PRIETA EARTHQUAKE ULF OBSERVATIONS

Of the mechanisms considered, electrokinetic effects generate the strongest magnetic field at the surface associated with a shear fracture driven by high pore pressures in response to failure of an impermeable layer between fault zone compartments. The 17 October 1989 Loma Prieta $M_L$ 7.1 earthquake offers a unique opportunity to test the validity of the fracture propagation model. Before and after this earthquake, an ultra-low frequency (ULF) magnetometer operated in Corralitos, California, only 7 km from the epicenter of the mainshock. This instrument indicated a number of ULF magnetic field fluctuations recorded as spectral peaks throughout its 0.01 to 10.0 Hz bandwidth (Fraser-Smith et al, 1990). Within the ULF range, the 0.01 to 0.02 Hz frequency window showed the most remarkable signals, with barely perceptible changes in the 5.0 to 10.0 Hz bandwidth.

The primary focus of this paper is determining whether the evolution of the compartment model proposed by Byerlee (1993) can generate magnetic signals large enough to account for the observations of Fraser-Smith et al (1990). Specifically, we wish to examine two major increases in the magnetic field of October 1989, one on 5 October 1989 and the other just three hours prior to the mainshock. In the 0.01 to 0.02 Hz bandwidth, the first had a magnitude of 2.0 nT and the second 6.7 nT.

The static considerations of the previous sections illustrate that electrokinetic effects generate signals comparable with the largest changes in the Fraser-Smith data. The question becomes how to explain the oscillatory nature of the signals. For the geometry of the fault zone described by Byerlee (1993), these oscillations can occur in two significant ways as mentioned in the fault model section. The first is a stop and start propagation of the fracture within the low pressure compartment. This mechanism is triggered because of the tremendous decrease in pore pressure from the rupture in the impermeable boundary between compartments to the fracture tip. This pressure decrease results in reduced solubility and deposition of silicates in the fracture. With enough deposition, the fracture may be sealed and propagation ceases. With the end of fluid motion, the electrokinetic effects also cease. The pressure in the newly sealed fracture rapidly rises to lithostatic levels, the new seal ruptures, and the process repeats. This process continues to repeat until the next lithostatic compartment is reached. Upon reaching the next lithostatic compartment, the deformation propagates at shear wave velocity to the next impermeable seal between high- and low-pressure compartments. The full process repeats itself again.

The length of the fracture determines the time it takes for the fracture to propagate that distance. If we assume that the fluid pressure rises virtually instantaneously to lithostatic levels once the fracture seals or is partially blocked by silicate deposition, then the fracture propagation time is equal to the period of oscillation for this intermediate stage. A relation between propagation time and fracture distance can be derived from the pore pressure distribution in the fracture shown in Figure 2. If we consider that deposition and initiation of the stop and start mechanism occurs under the same circumstances as those used for determination of fluid velocity in the shear fracture (see MHD section), we can find an expression which relates fracture length and time.

$$\frac{\ell}{2\sqrt{DT}} = 1.36 \quad or \quad T = \frac{\ell^2}{7.4D} \tag{15}$$

Since T is also the period of oscillation, the frequency window corresponding to a particular fracture length is simply 1/T.

The strain data and exhumed fault observations, along with determinations of average pore pressure in the fault zone, limit the compartment size and geometry. For equal width and height of the lithostatic and hydrostatic compartments, the hydrostatic compartments must be one-third the length of lithostatic compartments. Since the electrokinetic effects occur only in the low pressure compartments, the maximum fracture length would be about 300 hundred meters. For this case, we ignore stop and start fracture propagation in the low-pressure compartment and consider only the repetition as the fracture propagates along the fault. The period of oscillation is about 180 seconds, which corresponds to a frequency of 0.006 Hz. This frequency is just below

the ULF range in which signals were reported by Fraser-Smith et al (1990), and just above the frequency of strainmeters where no unusual activity was discovered in the days before the Loma Prieta earthquake (Johnston et al, 1990). Consequently, accepting the veracity of the Fraser-Smith data places an additional constraint on the compartment size vis-a-vis the fracture propagation rates.

Table 4 shows the fracture lengths which correspond to each frequency band in the ULF range. Consider the 0.5 Hz to 1.0 Hz bandwidth, with associated fractures of about 20 to 30 m in length. In the magnetic data recorded at Corralitos, the last obvious disturbances occur in this frequency window. Suppose that a 25 m fracture repeatedly forms within the low-pressure compartment until reaching the next high-pressure compartment. For a 200 m long low-pressure compartment, eight such fractures will form, generating a repeating magnetic signature. Figure 6 shows the magnetic field time series for an observer standing 17 km directly above the rupture between the high- and low-pressure compartments. Figure 7 shows the power spectrum which results from this repeated signal. The most obvious and important feature of the power spectrum is the increasing strength of the signal as frequency decreases. This matches the general character of the magnetic data recorded by Fraser-Smith et al (1990).

## CONCLUSIONS

This investigation of magnetic effects associated with the evolution of the compartment fault model proposed by Byerlee (1993) yields several interesting results. Most importantly, the rupture of brittle impermeable layers between compartments of different pore pressure regimes in response to tectonic loading leads to shear fracture formation in the low-pressure compartment. Extremely turbulent fluid flow into the shear fracture generates magnetic signals measurable at the surface as a result of electrokinetic effects. Other mechanisms, such as piezomagnetism and magnetohydrodynamics, contribute negligibly to the overall magnetic field.

Stop and start fracture propagation, a consequence of the spatial distribution of the pore pressure within the shear fracture, may explain the short period ULF signals recorded prior to the Loma Prieta earthquake. If this mechanism generated these measured signals, segments of the fault of perhaps several kilometers in extent were failing largely aseismically during this time. The cumulative effects of such large scale aseismic failure could have been detected on a strain array over the area had one existed at the time. Further, some increase in microseismicity should also have occurred, though none was apparent during the period prior to the Loma Prieta earthquake (White, 1993).

The aftershock record of the Loma Prieta earthquake shows some coseismic correlation between seismic events and magnetic field fluctuations (Fenoglio et al, 1993). Given that the compartment failure mechanism can trigger earthquakes, it is not unexpected that the aftershocks, occurring in an already weakened and deformed region, are coincident with magnetic field changes that might result from further fluid flow in response to continuing failure of the brittle impermeable layers between compartments.

Strain constraints and exhumed fault observations limit the size of the high-pressure compartments to less than 1 km in extent. By inference from the average pore pressure within the fault zone, the low pressure compartments are only one-third the extent of the high-pressure compartments. Even for low-pressure compartments only 200 m across, the electrokinetic effects at the surface are similar in magnitude to the recorded magnetic field data. This being the case, the compartment fault model appears capable of satisfying a number of observations associated with the Loma Prieta earthquake.

## REFERENCES

Andrews, D.J., Rupture Velocity of Plane Strain Shear Cracks, Journal of Geophysical Research, Vol 81, pp 5679-5687, 1976.

Barker, J.S., and D.H. Salzberg, Long-Period and Broad-Band Teleseismic Body-Wave Modeling of the October 18, 1989 Loma Prieta Earthquake, Geophysical Research Letters, Vol 17, pp 1409-1412, 1990.
Brace, W.F., Permeability of Crystalline and Argillaceous Rocks, International Journal of Rock Mechanics, Vol 17, pp 876-893, 1980.
Brace, W.F., J.B. Walsh, and W.T. Frangos, Permeability of Granite under High Pressure, Journal of Geophysical Research, Vol 73, pp 2225-2236, 1968.
Byerlee, J., Model for Episodic Flow of High Pressure Water in Fault Zones before Earthquakes, Geology, Vol 21, pp 303-306, 1993.
Byerlee, J., The Change in Orientation of Subsidiary Shears near Faults Containing High Pore Fluid Pressure, Tectonophysics, Vol. 211, pp 295-303, 1992.
Byerlee, J., Friction, Overpressure and Fault Normal Compression, Geophysical Research Letters, Vol 17, pp 2109-2112, 1990.
Chester, F.M., and J.M. Logan, Implications for Mechanical Properties of Brittle Faults from Observations of the Punchbowl Fault Zone, California, Pure and Applied Geophysics, Vol 124, pp 79-106, 1986.
Davis, P.M., D.R. Pierce, R.L. McPherron, D. Dzurisin, T. Murray, M.J.S. Johnston, and R. Mueller, A Volcanomagnetic Observation of Mount St. Helens, Geophysical Research Letters, Vol 11, pp 225-228, 1984.
Davis, P.M., F.D. Stacey, C.J. Zabock, and J.V. Olson, Improved Signal Discrimination in Tectonomagnetism; Discovery of a Volcanomagnetic Effects at Kilauea, Hawaii, Physics of the Earth and Planetary Interiors, Vol 19, pp 331-336, 1979.
Dobrovolsky, I.P., N.I. Gershenzon, and M.B. Gokhberg, Theory of Electrokinetic Effects Occuring at the Final Stage in the Preparation of a Tectonic Earthquake, Physics of the Earth and Planetary Interiors, Vol 57, pp 144-156, 1989.
Draganov, A.B., U.S. Inan, and Yu.N. Taranenko, ULF Magnetic Signatures at the Earth due to Ground Water Flow: a Possible Precursor to Earthquakes, Geophysical Research Letters, Vol 18 (6), pp 1127-1130, 1991.
Eberhart-Phillips, D., V.F. Labson, W.D. Stanley, A.J. Michael, and B.D. Rodriguez, Preliminary Velocity and Resistivity Models of the Loma Prieta Earthquake Region, Geophysical Research Letters, Vol 17, pp 1235-1238, 1990.
Fenoglio, M.A., A.C. Fraser-Smith, G.C. Beroza, and M.J.S. Johnston, Comparison of Ultra-Low Frequency Electromagnetic Signals with Aftershock Activity during the 1989 Loma Prieta Earthquake Sequence, Bull. Seis. Soc. Amer., Vol 83, pp 347-357, 1993.
Fitterman, D.V., Theory of Electrokinetic-Magnetic Anomalies in a Faulted Half-Space, Journal of Geophysical Research, Vol 84, pp 6031-6040, 1979.
Fitterman, D.V., Electrokinetic and Magnetic Anomalies Associated with Dilatant Regions in a Layered Earth, Journal of Geophysical Research, Vol 83, pp 5923-5928, 1978.
Fitterman, D.V., and T.R. Madden, Resistivity Observations During Creep Events at Melendy Ranch, California, Journal of Geophysical Research, Vol 82, pp 5401-5408, 1977.
Fournier, R.O., and R.W. Potter II, An Equation Correlating the Solubility of Quartz in Water from 25 Degrees to 900 Degrees C at Pressures up to 10,000 Bars, Geochemica et Cosmochimica Acta, Vol 46, pp 1969-1973, 1982.
Fournier, R.O., Continental Scientific Drilling to Investigate Brine Evolution and Fluid Circulation in Active Hydrothermal Systems, IN: Observation of the Continental Crust through Drilling I, ed. C.B. Raleigh, pp 98-122, 1985.
Fraser-Smith, A.C., A. Bernardi, P.R. McGill, M.E. Ladd, R.A. Helliwell, and O.G. Villard, Jr., Low-Frequency Magnetic Field Measurements near the Epicenter of the $M_S$ 7.1 Loma Prieta Earthquake, Geophysical Research Letters, Vol 17 (9), pp 1465-1468, 1990.
Herrington, R.J., and J.J. Wilkinson, Colloidal Gold and Silica in Mesothermal Vein Systems, Geology, Vol 21, pp 539-542, 1993.
Ishido, T., and H. Mizutani, Experimental and Theoretical Basis of Electrokinetic Phenomena in Rock-Water Systems and its Application to Geophysics, Journal of Geophysical Research, Vol 86, pp 1763-1775, 1981.

Johnston, M.J.S., A.T. Linde, and M.T. Gladwin, Near-Field High Resolution Strain Measurements Prior to the October 18, 1989, Loma Prieta $M_S$ 7.1 Earthquake, Geophysical Research Letters, Vol 17, pp 1777-1780, 1990.

Johnston, M.J.S., Review of Magnetic and Electric Field Effects near Active Faults and Volcanoes in the U.S.A., Physics of the Earth and Planetary Interiors, Vol 57, pp 47-63, 1989.

Johnston, M.J.S., and R.J. Mueller, Seismomagnetic Observation during the 8 July 1986 Magnitude 5.9 North Palm Springs Earthquake, Science, Vol 237, pp 1201-1203, 1987.

Kerrich, R., W.S. Fyfe, and I. Allison, Iron Reduction around Gold-Quartz Veins, Yellowknife District, Northwest Territories, Canada, Economic Geology, Vol 72, 1977.

Madden, T.R., Electrical Measurements as Stress-Strain Monitors, U.S.G.S, Open File Report (79-0370), pp 301-347, 1979.

Mazella, A., and H.F. Morrison, Electrical Resistivity Variations Associated with Earthquakes on the San Andreas Fault, Science, Vol 185, pp 855-857, 1974.

Miyakoshi, J., Anomalous Time Variation of the Self-Potential in the Fractured Zone of an Active Fault Preceding the Earthquake Occurence, J. Geomag. Geoelect., Vol 38, pp 1015-1030, 1986.

Mizutani, H., T. Ishido, T. Yokokura, and S. Ohnishi, Electrokinetic Phenomena Associated with Earthquakes, Geophysical Research Letters, Vol 13, pp 365-368, 1976.

Morrison, H.F., R. Fernandez, and R.F. Corwin, Earth Resistivity, Self Potential Variations, and Earthquakes: a Negative Result for M=4.0, Geophysical Research Letters, Vol 6, pp 139-142, 1979.

Nadeau, R., M. Antilok, P. Johnson, W. Foxall, and T.V. McEvilly, Seismological studies of Parkfield III: Microearthquake clusters in the study of fault-zone dynamics, Bull. Seis. Soc. Am. (in press), 1993.

Nourbehecht, B., Irreversible Thermodynamic Effects in Inhomogeneous Media and Their Applications in Certain Geoelectric Problems, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, 1963.

Overbeek, J.Th., Thermodynamics of Electrokinetic Phenomena, Journal of Colloid Science, Vol 8, p 420, 1953.

Park, S.K., and D.V. Fitterman, Sensitivity of the Telluric Monitoring Array in Parkfield, California to Changes of Resistivity, Journal of Geophysical Research, Vol 95, pp 15557-15571, 1990.

Probstein, R.F., and R.E. Hicks, Removal of Contaminants from Soils by Electric Fields, Science, Vol 260, pp 498-503, 1993.

Qian, F., Y. Zhao, M. Zu, Z. Wang, X. Liu, and S. Chang, Geoelectric Resistivity Anomalies before Earthquakes, Scientia Sinica (Series B), Vol 26, pp 326-336, 1983.

Rice, J.R., Fault Stress States, Pore Pressure Distributions, and the Weakness of the San andreas Fault, in Evans, B., and Wong, T.F., Eds., Earthquake Mechanics and Transport Properties of Rocks, Academic Press, London, pp 475-503, 1992.

Sasai, Y., Tectonomagnetic Modeling on the Basis of the Linear Piezomagnetic Effect, Bull. Earthq. Res. Inst., Univ. Tokyo, Vol 66, pp 585-722, 1991.

Sasai, Y., Application of the Elasticity Theory of Dislocation to Tectonomagnetic Modeling, Earthquake Research Institute Bulletin, Vol 55, pp 387-447, 1980.

Sheng, P., and Z. Chen, Local-Field Distribution in Random Dielectric Media, Phys. Rev. Lett., Vol 60, P 2277, 1988.

Shercliff, J.A., The Theory of Electromagnetic Flow-Measurement, University Press, Campbridge, 146 pp, 1962.

Sibson, R.H., F. Robert, and K.H. Poulsen, High-angle Reverse Faults, Fluid-induced Pressure Cycling, and Mesothermal Gold-Quartz Deposits, Geology, Vol 16, pp 551-555, 1988.

Stacey, F.D., and M.J.S. Johnston, Theory of the Piezomagnetic Effect in Titanomagnetic-Bearing Rocks, Pure and Applied Geophysics, Vol 97, pp 146-155, 1972.

Stacey, F.D., the Seismomagnetic Effect, Pure and Applied Geophysics, Vol 58, 1964.

White, R.A., The Loma Prieta, California Earthquake of October 17, 1989 - Preseismic Observations, USGS Professional Paper 1550-C, 1993.

Table 1. Constants important for electrokinetic effects

| Variable | Value |
|---|---|
| $\varepsilon$, dielectric constant | $7.07 \times 10^{-10}$ F/m |
| $\zeta$, zeta potential | 0.1 V |
| $\beta$, fluid compressibility | $3 \times 10^{-10}$ Pa$^{-1}$ |
| $\varphi$, porosity | 0.5 |
| $\eta$, dynamic viscosity | $10^{-4}$ Pa s |
| $\mu_o$, magnetic permeability of free space | $4\pi \times 10^{-7}$ H/m |
| k, fracture permeability | $10^{-12}$ m$^2$ |
| $\rho$, density of rock | 2500 kg/m$^3$ |
| $\ell$, fracture length | 200 m |
| d, depth to the top of the fracture | 17 km |
| h, height of fracture | 1000 m |

Table 2. Constants important for magnetohydrodynamic effects

| Variable | Value |
|---|---|
| $\mu_o$, magnetic permeability of free space | $4\pi \times 10^{-7}$ H/m |
| s, fluid conductivity | 0.1 S/m |
| $\beta$, fluid compressibility | $3 \times 10^{-10}$ Pa$^{-1}$ |
| $\varphi$, porosity | 0.5 |
| $\eta$, dynamic viscosity | $10^{-4}$ Pa s |
| $B_o$, Earth's magnetic field strength | $5 \times 10^{-5}$ T |
| k, fracture permeability | $10^{-12}$ m$^2$ |
| $\ell$, fracture length | 200 m |

Table 3. Constants important for piezomagnetic effects

| Variable | Value |
|---|---|
| S, stress sensitivity | $1.5 \times 10^{-4}$ bar$^{-1}$ |
| $t_{jo}$, average magnetization | $3 \times 10^{-3}$ emu/cc |
| G, shear modulus | $3 \times 10^{10}$ Pa |
| d, dislocation | 8.33 m |
| H, Curie depth | 25 km |
| $\theta$, magnetic dip of ambient field | 62° |
| $\psi$, magnetic declination | N17°E |
| h, fracture height | 1000 m |
| d, depth to the top of the fracture | 17 km |
| $\ell$, fracture length | 200 m |

Table 4. Relationship of frequency window and fracture length

| Frequency (Hz) | Period (sec) | Fracture Length (m) |
|---|---|---|
| 0.01 to 0.02 | 50 to 100 | 157 to 222 |
| 0.02 to 0.05 | 20 to 50 | 99 to 157 |
| 0.05 to 0.10 | 10 to 20 | 70 to 99 |
| 0.1 to 0.2 | 5 to 10 | 50 to 70 |
| 0.2 to 0.5 | 2 to 5 | 31 to 43 |
| 0.5 to 1.0 | 1 to 2 | 22 to 31 |
| 1.0 to 2.0 | 0.5 to 1 | 16 to 22 |
| 2.0 to 5.0 | 0.2 to 0.5 | 9.9 to 16 |
| 5.0 to 10.0 | 0.1 to 0.2 | 7.0 to 9.9 |

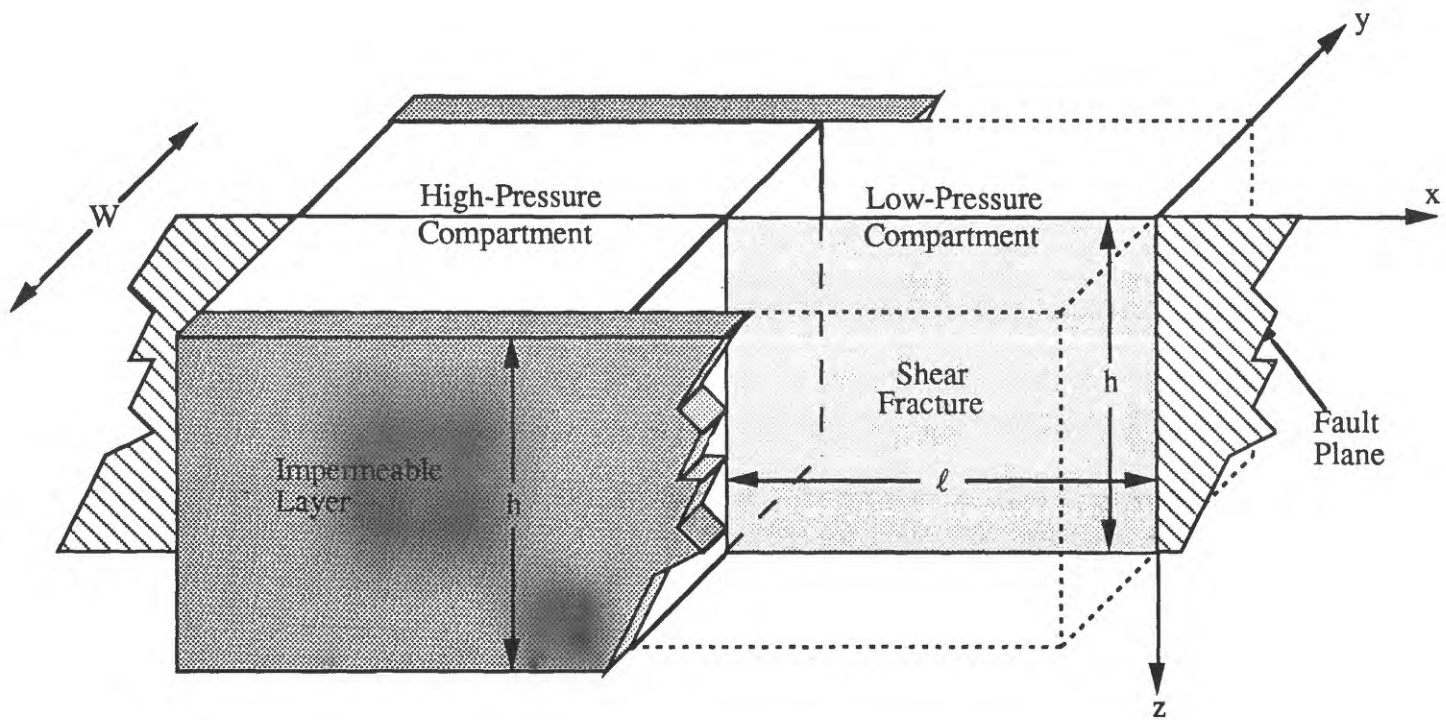



Figure 1. Diagram of the essential components of the Byerlee compartment model. The impermeable layers (the darkly shaded regions) prevent fluid communication between the fault zone and the surrounding country rock. The shear fracture forms along the fault plane as a result of a rupture in the impermeable seal between the high- and low-pressure compartments. W is the width of the fault zone, h is the height of the high-pressure compartment (and the shear fracture), and $\ell$ is the length of the shear fracture.

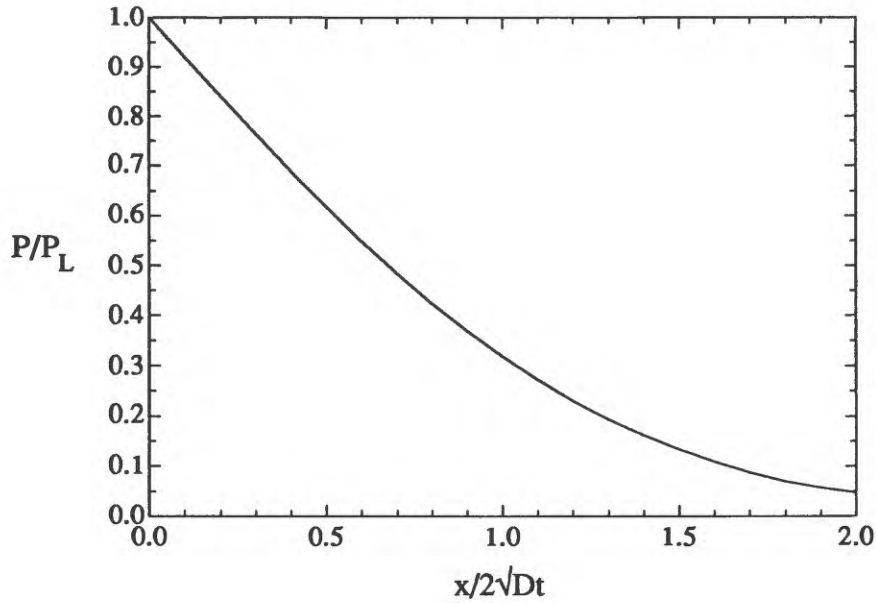



Figure 2. This plot shows the approximate normalized pore pressure distribution in the shear fracture as a function of distance and time. 95% of the total distribution occurs in the region $0 < x/2\sqrt{Dt} < 1.36$.

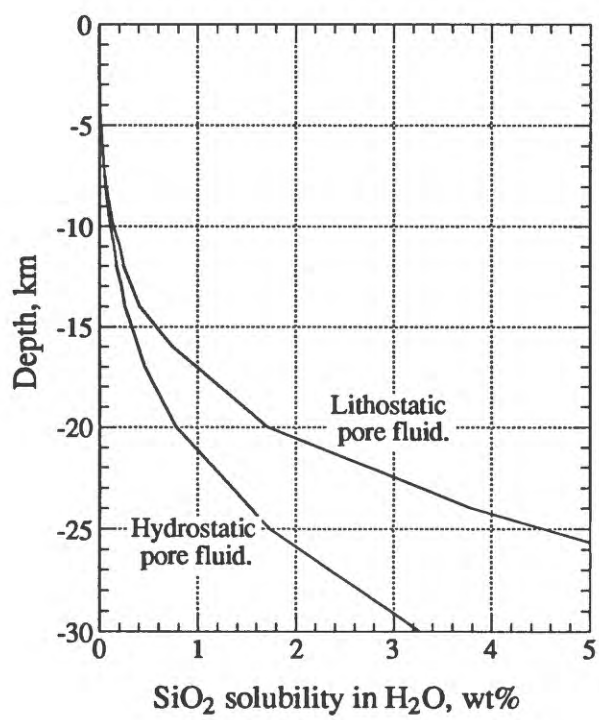



Figure 3. This figure illustrates the solubility of silica in water by weight precentage versus depth on the San Andreas Fault for hydrostatic and lithostatic pore pressure gradients (from Fournier and Potter, 1982). The model assumes a geotherm for a non-self-heating (weak) fault (C. Williams, A. Lachenbruch, and J. Sass, unpublished data).

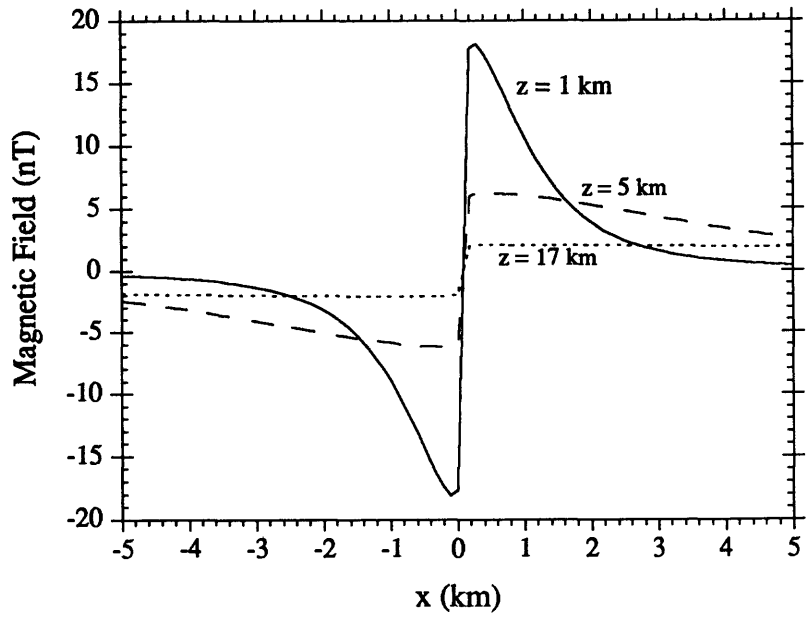



Figure 4. This plot compares the horizontal magnetic field along fault strike generated by fluid flow in a shear fracture 200 m long and 1000 m high buried at 1 km, 5 km, and 17 km (at time t = 1 second).

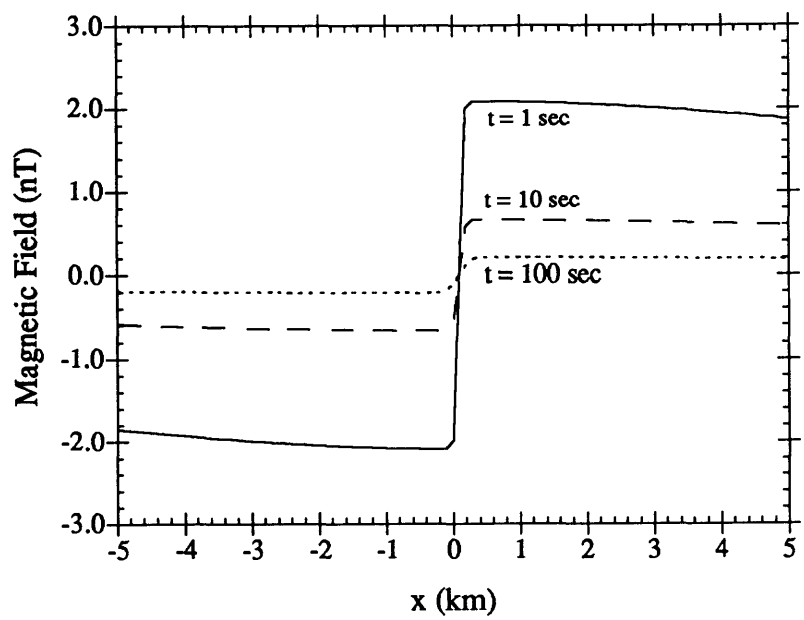



Figure 5. This plot shows the horizontal magnetic field evolution along fault strike generated by fluid flow in a shear fracture 200 m long and 1000 m high buried at 17 km. The evolution is showed at times t = 1, 10, and 100 seconds.

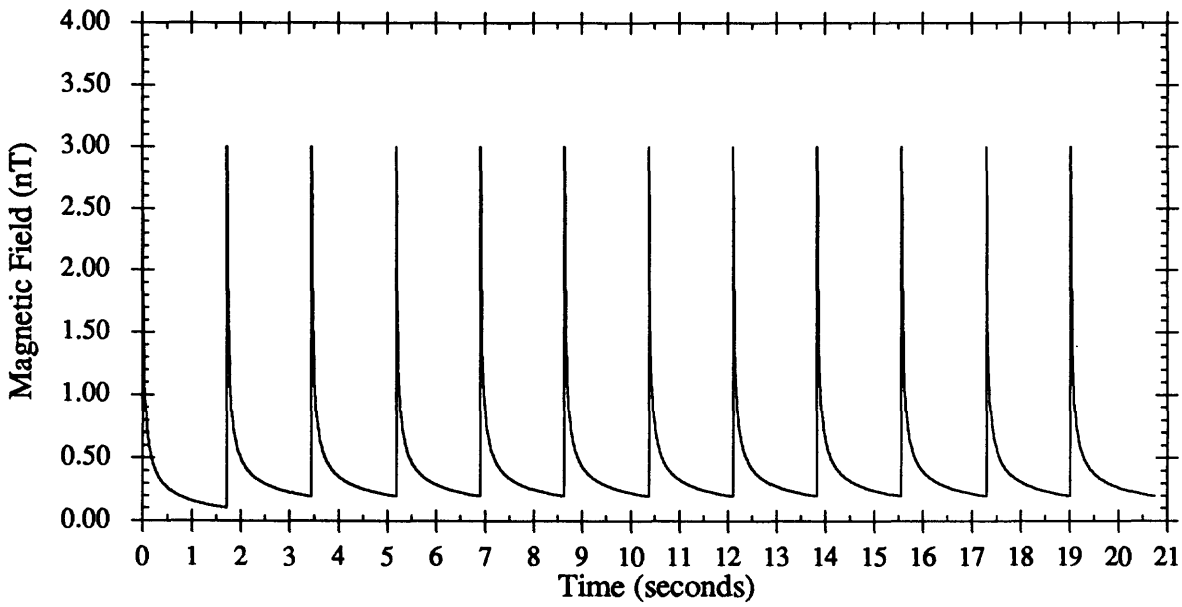



Figure 6. This graph shows the horizontal magnetic field for 25 m fractures as seen by an observer at the initial rupture, 17 km above the source. The source is fluid flowing into a 200 m by 1000 m shear fracture.

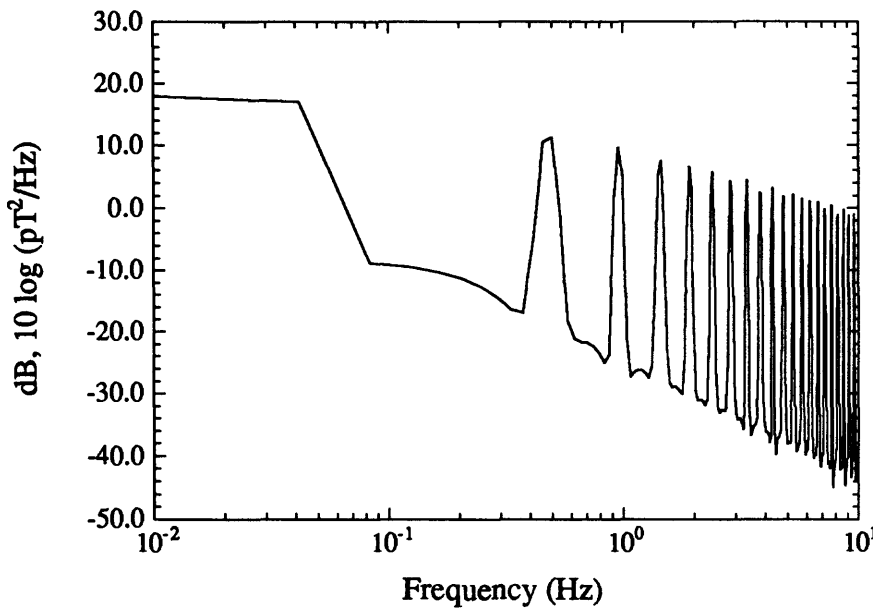



Figure 7. This is the power spectrum for the time series in Figure 6. The general characteristics of this plot are similar to the energy distribution in the magnetic data recorded by Fraser-Smith et al (1990).

# PRESSURE SOLUTION AS A MECHANISM FOR CRACK SEALING AROUND FAULTS

## Natural and Experimental Evidence

J.P. Gratier [1,2], T. Chen [1], R. Hellmann [1]

[1]*LGIT-CNRS (URA 733) Université Joseph Fourier, IRIGM, BP 53X, 38041 Grenoble (France)*
[2]*Institute for Crustal Studies, University of California, Santa Barbara, CA 93106*

## INTRODUCTION

Several models pertaining to earthquake cycles imply intermittent fluid flow through fault zones (Sibson, 1977; Lachenbruch, 1980; Byerlee, 1990; Rice, 1992; Sleep & Blanpied, 1992). Fluid pressure in a fault zone is expected to change from a high value before an earthquake (leading to a progressive weakening of the fault) to a low value after the event (due to dilatancy associated with the rupture). To account for the low frictional heating and lack of strength of active faults, as inferred from low heat flow measurements (Hickman, 1991), earthquakes are postulated to occur when fluid pressures are close to lithostatic levels. Evidence for high fluid pressures has been found in tectonically active areas. As an example, in the Alps, a comparison between fluid density in horizontal veins (from fluid inclusion studies), temperature (from K/Na ratios), and the thickness of the cover, showed that fluid pressure in those veins was very near lithostatic values at depths ranging from 6 to 10 km (Poty et al., 1974; Bernard et al., 1977). Based on borehole records from gas and oil wells, Hubert & Rubey (1959) and Berry (1973) showed many examples of pore pressures largely above hydrostatic pressures, sometimes even approaching lithostatic pressures at depths ranging from 0.3 to 4 km.

In seismic zones, earthquakes are expected to change the permeability of the country rock and may temporarily reduce fluid pressures to hydrostatic values (Sibson, 1992; Chester et al. 1993; Boullier & Robert, 1992). During the interseismic period, the progressive increase of fluid pressure from near hydrostatic to near lithostatic levels is therefore a crucial parameter which may play an important role in the mechanisms leading to earthquakes (Sibson, 1977; Lachenbruch, 1980; Byerlee, 1990; Rice, 1992; Sleep & Blanpied, 1992). To achieve such high fluid pressures in faults zones, different geodynamic mechanisms have been proposed, including flow of deep fluids at high pressure (Rice, 1992) and compaction within fault gouges (Sleep & Blanpied, 1992). In these models, the very rapid sealing of cracks is postulated to occur in the country rock after the occurrence of an earthquake. The subsequent increase in the fluid pressure is presumed to be only dependent on the rates of these geodynamic processes. This assumption of rapid sealing is based upon experimental results of crack healing (Smith & Evans, 1984; Brantley et al., 1990). However, self-healing may not necessarily be the dominant mechanism associated with natural crack sealing. If the rate of change of permeability in the country rock is not fast relative to the geodynamic rates, then changes in the permeability may be the key factor in controlling the rate of increase in fluid pressure during the interseismic period.

The aim of our study was, therefore, to consider the different mechanisms and the kinetics of crack sealing near active faults. Our study used two approaches: natural observations were made to identify the mechanisms of crack sealing, and secondly, pressure solution experiments allowed

us to investigate the kinetics of the sealing process. The main result of this study is the finding that natural sealing is usually not achieved by a self-healing process. Pressure solution, which is based on the transfer of material from solution cleavages to cracks, appears to be a much more important mechanism for crack sealing. Based on our experimental results, the kinetics of pressure solution are estimated to be rather slow, similar to the recurrence time of earthquakes. The kinetics of crack sealing by pressure solution may thus be important in controlling the increase in fluid pressure around faults, as well as the mechanisms associated with critical failure.

## CRACK SEALING MECHANISMS

*Mechanical compression* can lead to crack closure (Brace, 1972); however, as cracks often open perpendicular to the direction of least effective stress, their closure implies that the effective stress directions have changed either by a decrease in the fluid pressure, or by a change in the state of stress; however, variations in stress directions will not be considered here. The mechanical closure of microcracks may be indicated by the presence of continuous (non-sealed) microcracks. Another mechanism, *plastic flow*, can also result in the reduction of pore volume within a matrix. However, this flow mechanism should be important only at higher temperatures and pressures, such as below the seismic/aseismic transition zone (Kirby, 1983).

*Chemical processes* can also lead to crack sealing, where different mechanisms may be operative, depending on the aperture of the cracks (Wilkins et al., 1985). For the case of low aperture cracks (0.3 - 3 μm; Brantley et al., 1990), *self-crack healing* is driven by the lowering of the crack surface free energy (Smith & Evans, 1984). This mechanism may be evidenced by trails of fluid inclusions along the healed crack surfaces (Fig. 1a). On the other hand, cracks may also eventually become self-sealed without any traces of fluid inclusions (B. Evans et al., this volume). Large aperture cracks (10 to 200 μm) generally show evidence for the input of material, as is commonly found in naturally-occurring crack-seal structures (Ramsay, 1980; Cox & Etheridge, 1983; Gratier & Gamond, 1990; Boullier & Robert, 1992). Various mechanisms, which are not always easily distinguishable, can control this process. Processes which can lead to the transfer and deposition of material in a crack are discussed in more detail below.

A change in *pressure* generally has only a small effect on mineral solubility. An exception to this is the case of fluid boiling (Hedenquist et al., 1992), which may be induced by a pressure drop associated with earthquakes (Sibson, 1987). This mechanism for crack sealing is associated with mass transfer by infiltration. When considering the irregular nature of paths associated with mass transfer, only the narrowest parts would be sealed after fluid infiltration and deposition had occurred (Fig. 1b). Since boiling would be associated with rapid fluid displacement, asperities along the fault might induce local drops in the fluid pressure and promote localized deposition in pressure shadow zones. Asymmetric dissolution precipitation near asperities (Fig. 1b') may, therefore, be evidence for this mechanism and indicate the sense of fluid displacement. This type of mechanism has also been evidenced in experimental studies (Lefaucheux, 1988).

*Temperature* changes can lead to the sealing of large fractures based on a change in mineral solubility. However, this effect is not the same for all minerals. As an example, quartz and calcite display regular and inverse solubility versus temperature dependencies, respectively (Fyfe et al., 1978). For the case of a normal solubility-temperature dependence, a progressive decrease in temperature leads to widespread deposition (Cathles, 1977). The sealing of cracks with this mechanism is also based on the infiltration of fluids into the matrix. However, complete sealing is probably not possible in that after the sealing of the narrowest paths, some isolated void spaces still remain open (Fig. 1c). *Fluid composition and pH* changes can also lead to either deposition (fluid oversaturation) or dissolution (fluid undersaturation) in open veins (Brimhall & Crerar,

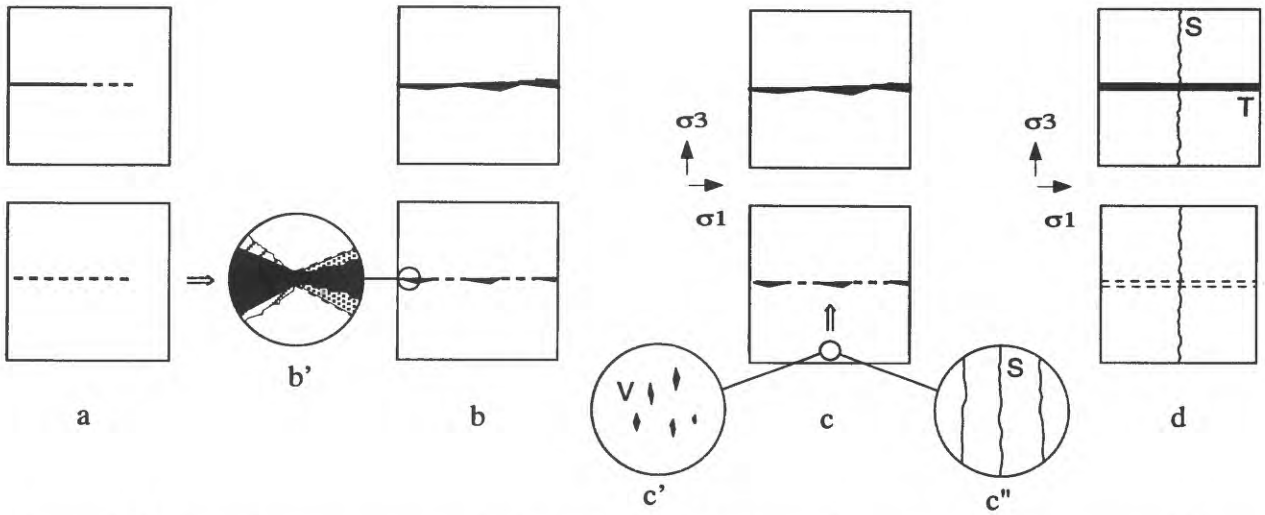



Figure 1: Microstructures expected with various mechanisms of crack sealing, geometry before (top) and after (bottom) sealing: solid is white, fluid is black. (a) Trails of fluid inclusions along the healed crack surfaces with self-healing (no external input of material); (b) and (c) Input of material associated with fluid infiltration, only the narrowest part of the path was sealed; (b') with rapid fluid displacement (double arrow) dissolution occurs on surface subjected to high stress (hatched) and precipitation occurs on surface with low stress (shaded) around the same asperity. With fluid infiltration (double arrow) the kinetics of sealing may be reaction rate controlled (c') with free-face dissolution around voids in matrix (V), or diffusion rate controlled (c") with water-film diffusion along solution cleavage surfaces (S). (d) Complete sealing with mass transfer by diffusion from solution cleavage (S) to microcrack (T); two trails of fluid inclusion mark the initial limits of the crack.

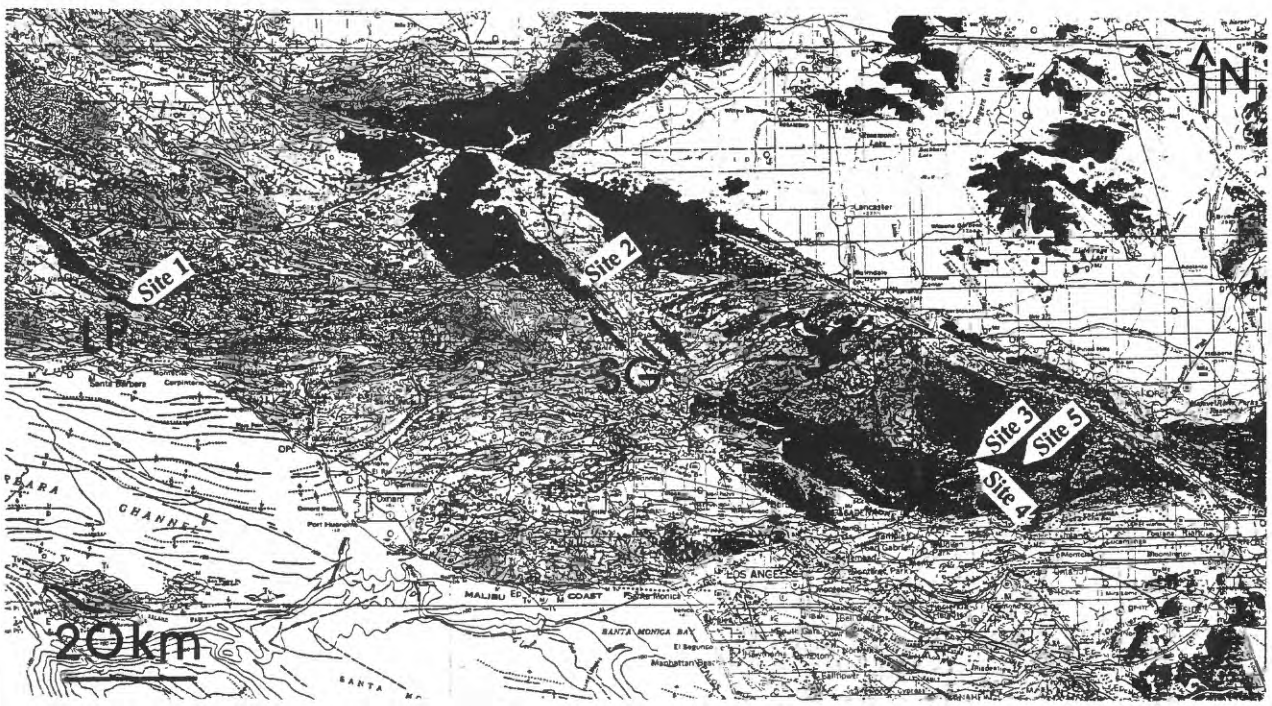



Figure 2: Geological map of the Transverse Range Province (from Jennings, 1977) and location of the sampling sites: site 1 - Little Pine fault (LP) and sites 2,3,4, & 5 - San Gabriel fault (SG).

1987; Wintsh et al, this volume). Special conditions would therefore be required to produce systematic crack sealing.

*Pressure solution-deposition* creep and compaction occur during diagenesis and tectonic processes. Under a deviatoric stress, soluble species are removed from stylolites or solution cleavage and redeposited in veins or voids (Weyl, 1959; Arthaud & Mattauer, 1969; Rutter, 1983). This mechanism has been observed in cataclastic fault zones (Sprunt & Nur, 1977; Shimamoto et al., 1991; Lloyd & Knipe, 1992) and can lead to the complete sealing of cracks (Fig. 1d) due to mass transfer by diffusion at a micrograin scale.

Evidence of fluid infiltration coupled with diffusion have been reported by several authors (Etheridge et al., 1984; McCraig & Knipe, 1990; Gratier & Gamond, 1990; Marquer & Burkhard, 1992). The strain rate associated with pressure solution is likely to be dependent on the kinetics of several successive processes: the kinetics of dissolution or precipitation reactions at the solid/fluid interface, or the rate of mass transfer by either diffusion or infiltration. Given that these processes occur in series, the slowest process becomes the rate-limiting step (Raj, 1982; Rutter, 1983). Pressure solution of grains under deviatoric stress leads to at least two types (Fig. 1c', 1c") of mineral/fluid interactions. When mass transfer occurs by fluid infiltration coupled with diffusion, the kinetics of the sealing process is strongly dependent on the mechanism of the fluid/mineral interactions. When dissolution occurs on the free faces of grains, diffusion coefficients are high (for example: $10^{-9}$ $m^2/s$ for silica at 350° (Ildefonse 1980)) and the limiting step becomes the rate of detachment reactions at the solid/fluid interface (Gratier & Jenatton, 1984; Tada & Siever, 1986). In this case the kinetics of sealing may be relatively fast. Evidence for free-face dissolution consists of pores, voids, and etch pits in the matrix (V, Fig. 1c'). The other possibility is that dissolution occurs within a trapped fluid film between grain contacts. Since mass transfer by diffusion is *much slower* than in a free fluid (for example, the diffusion coefficient for silica is $10^{-13}$ $m^2/s$ for water-film diffusion at 350°C; Gratier and Guiguet, 1986), the limiting step is not the rate of reaction at the solid/fluid interface, but rather the rate of mass transfer by water-film diffusion (Rutter, 1983; Gratier & Guiguet, 1986; Spiers & Schutjens, 1990). In this case the kinetics of sealing is always slow (see discussion). Evidence for this mechanism is the presence of solution cleavage seams (S, Fig. 1c").

## NATURAL CRACK SEAL FEATURES

Samples were collected from 5 sites near two major faults in California (Little Pine fault and San Gabriel fault) that were recently uplifted and exposed. This permitted the observation of the crack seal process in rocks brought up from depths of several kilometers. The locations of the sites are given in Fig. 2.

### Little Pine Fault

Examples of dissolution features associated with the progressive opening and deformation of tension gashes have been found in Upper Miocene Monterey shales near the Little Pine fault (site 1, Fig. 2). Solution cleavage parallel to the axial plane of large tight folds (several hundreds meters in size) is found kilometers away from the major fault. Folds axes and cleavage are sub-parallel to the fault and attest to intense shortening more or less perpendicular to this fault.

Samples a few meters away from the fault show nearly vertical tension gashes that register the progressive deformation of the rocks (Fig. 3a): younger veins (Tl) are almost undeformed and are perpendicular to the nearly vertical solution cleavage surfaces (S); earlier veins (Tf) are tightly folded and locally dissolved near large zones of solution cleavage; intermediate fold geometries may be found as well, indicating a progressive shortening during the formation of the veins. This

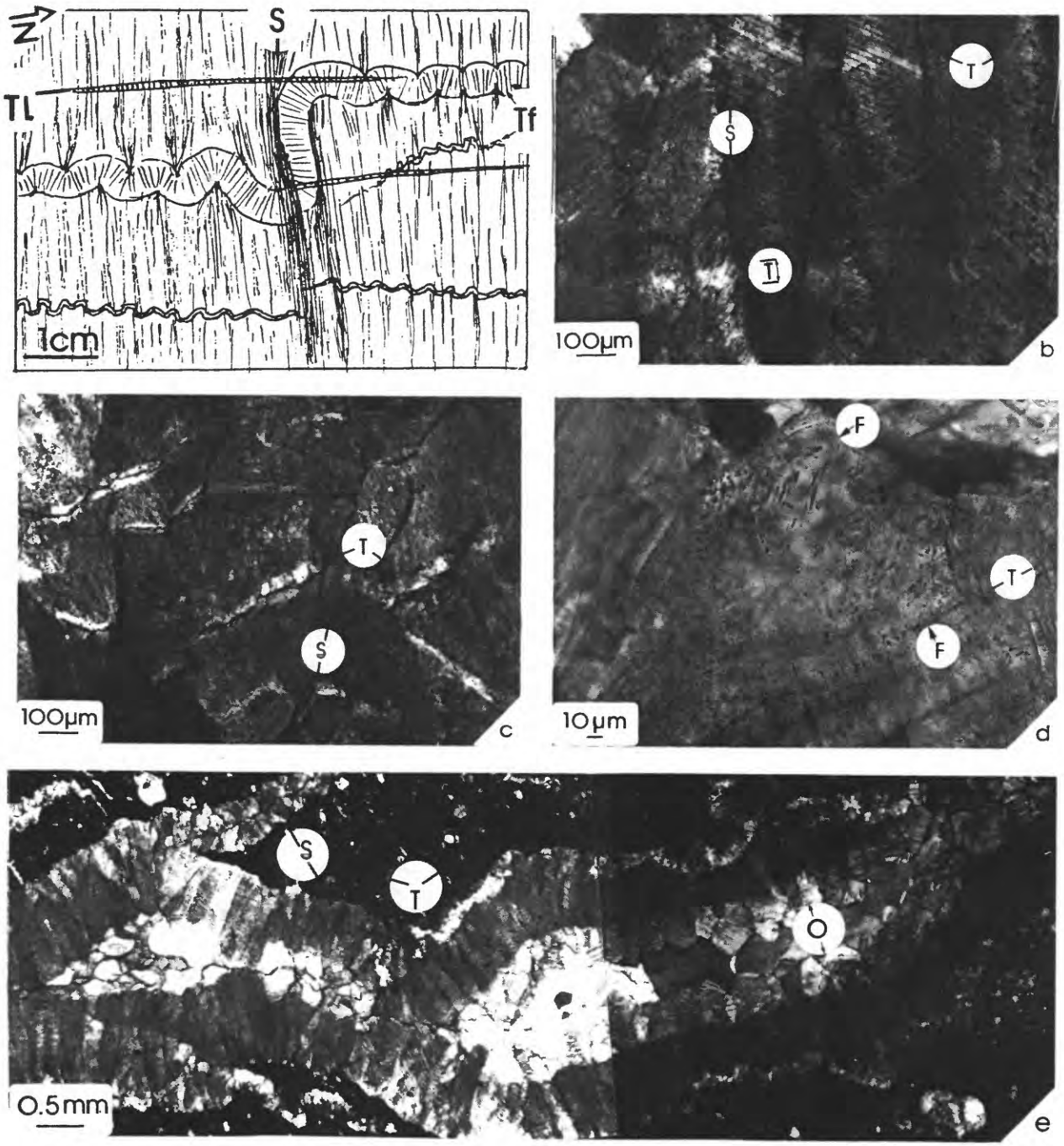



Figure 3: Solution cleavage (S) and crack sealing (T) (with input of material), near the Little Pine fault (site 1, Fig. 2). (a) Sketch of an outcrop, the younger veins (Tl) are almost undeformed and perpendicular to the solution cleavage (S), earlier veins (Tf) are tightly folded and locally dissolved. (b) & (c) Crack seal mechanism in calcite veins, each crack event either affects the whole vein (b) or is localized to just the outer limits of some grains (c). (d) Trail of fluid inclusions (F) which may indicate local self-crack healing. (e) In some cases, the last opening of some veins was large enough to break the continuity of the crack sealing process, leading to a change in the nature of deposition (O).

shortening perpendicular to the cleavage may be estimated by the unfolding of the later veins. The shortening reaches a minimum mean value of 30% for the outcrops (ten-hundred meters). The elongation value indicated by the veins is only about 1-2% (always the mean value for the outcrop). Precipitation might have occurred on a grain scale, thereby reducing the porosity of the rocks. However, no evidence of this is found in thin section. The amount of dissolution thus appears to have been much greater than the amount of precipitation. The extent of the closed system may have encompassed more than just the zone containing the solution cleavage. In such a case, mass transfer occurred by infiltration and diffusion-limited pressure solution, resulting in the removal of material from the shales.

Some veins are sealed by synaxial calcite fibers, indicating successive crack sealing under a nearly constant state of stress. Each crack seal event (T) either affects the whole vein (Fig. 3b), or is localized just at the outer limit of some grains (Fig. 3c). The large aperture of each successively sealed crack (from 10 to 100 μm, Fig. 3b, c, e), which is associated with the input of material, is incompatible exclusively with a self-healing process, even though markers for such a process may be found (trend of fluid inclusions, F, Fig. 3d). In some cases, the last opening of some veins was large enough to break the continuity of the crack seal process, leading to a change in the style of deposition. This is indicative of large permanent openings, which may have been associated with a high degree of fluid infiltration (O, Fig. 3e).

In conclusion, these observations suggest that the sealing of cracks, due to the input of material, is associated with pressure solution seams (S), both around faults (solution cleavage, Fig. 3a) and within veins (lateral dissolution of the fibers, Fig. 3b,c,e). Mass transfer is postulated to have occurred by fluid infiltration coupled with diffusion-limited pressure solution.

## San Gabriel Fault

Samples were collected from 3 sites several meters to hundreds of meters away from the central gouge zone of the San Gabriel fault (in granites and metamorphic rocks; sites 3,4,5, Fig. 2). In addition, one sample was taken directly from within the San Gabriel fault gouge (site 2, Fig. 2). Sketches of thin sections (Fig. 4) give the main displacement and the principal deformation features. Three of the thin sections were cut in an horizontal plane (sites 2,3,5), perpendicular to several meter-long strike slip faults with displacements on the order of a centimeter. One sample was cut in a vertical plane normal to an E-W trending cleavage (site 4).

The fault displacement was accommodated by a typical cataclastic process (Fig. 5a). Secondary fractures near the narrowest cataclastic zone were developed either as Riedel fractures (R, Fig. 4 & 5a, site 2 in the main gouge) or as conjugate fractures (C & C', Fig. 4, sites 3 & 5 in country rock). Away from the cataclastic zones (Fig. 4, 5), the crack orientation (T) indicates the direction of shortening and ranges from 20° to 45° with respect to the faults.This variation may be related to a different state of stress between the country rock and the main gouge zone, but it may also have been associated with a local reorientation of the stress state within bridges connecting segments of faults, or to the rotation/distortion of some blocks. Several dissolution and sealing processes can be distinguished in the samples.

Away from the cataclastic zones, the sealed cracks (T) are associated with and are perpendicular to solution cleavage (S), Fig. 5. This solution cleavage can be seen at the scale of several grains (stylolites with partial dissolution of some minerals, Fig. 5c) or at the scale of a single grain boundary (offset of dissolved markers along a stylolitic surface, Fig. 5d,e). Quartz and feldspar both show evidence (see Fig. 5) for dissolution (solution cleavage, S) and precipitation (crack sealing, T). The aperture of the cracks (or at least the thickness of the filling material) is easy to estimate when the sealing mineral is different from the host mineral (Fig. 5f,g,h,i). Crack apertures range from 10 to 100 μm. Apertures are more difficult to estimate when the sealing and the host mineral are not easily differentiated. An example of transgranular cracks (T, Fig. 5i)

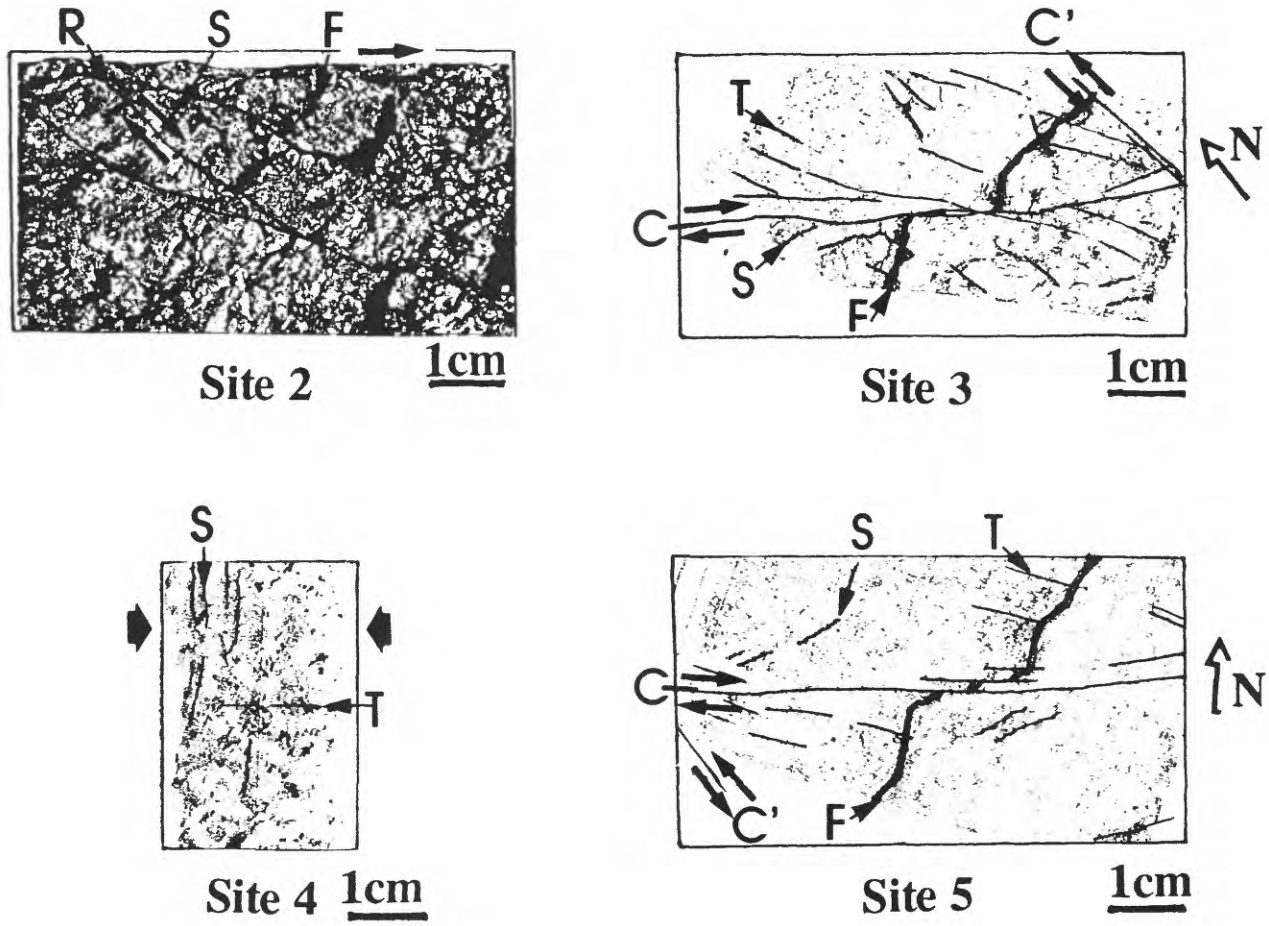



Figure 4: Thin sections from samples collected from sites 2,3,4, & 5 (Fig. 2) in granites and metamorphic rocks near the San Gabriel fault. Fault with gouge (C and C'), Riedel fracture (R), foliation (F), solution cleavage (S), crack (T). Sections in horizontal planes for sites 2,3, & 5; arrows indicate North, and in vertical E-W trending plane for site 4.

Figure 5: Photographs showing microstructures located at different sites some distance away from the faults: (a) & (b) = site 2, (c) & (d) = site 5, (e) & (f) = site 3, (g) & (f) = site 4. Arrows indicate the orientation and relative displacement of the nearest fault with gouge (C or C'); Riedel fracture (R), solution cleavage (S), crack (T) and growth of micas along the main infiltration paths (M).

Figure 6: Photographs showing microstructures within the gouge of the faults in different sites: (a), (b), (d), (e), (f) & (g) = site 5; (c) & (h) = site 2; Arrows indicate the orientation and relative displacement of the nearest fault with gouge (C or C'); solution cleavage (S), indentation by dissolution (I) and zone of local plasticity (P).

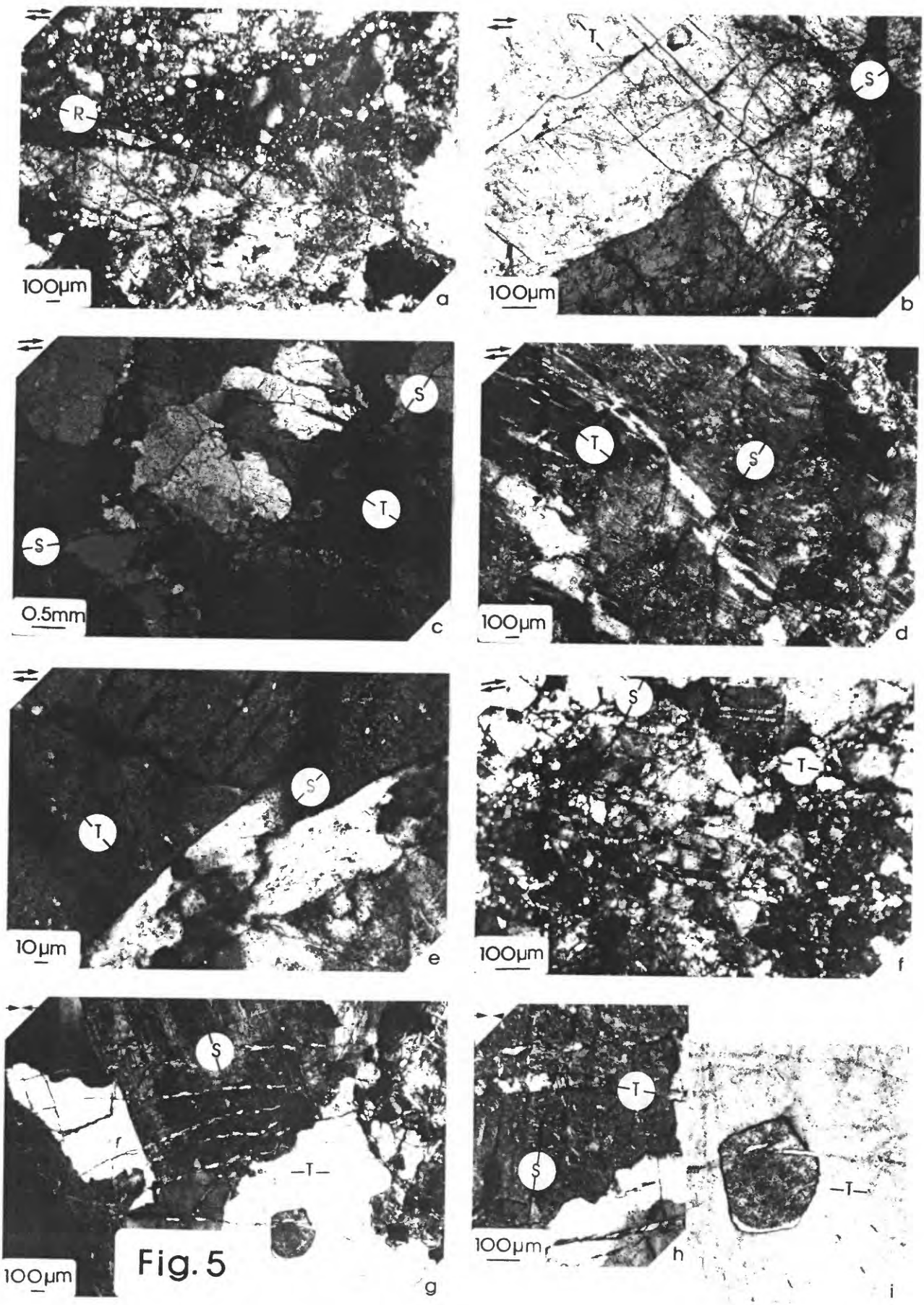

R
100μm
a
T
S
100μm
b
S
T
S
0.5mm
c
T
S
100μm
d
T
S
10μm
e
S
T
100μm
f
S
r
T
100μm
g
T
S
100μm
h
T
i

Fig. 5

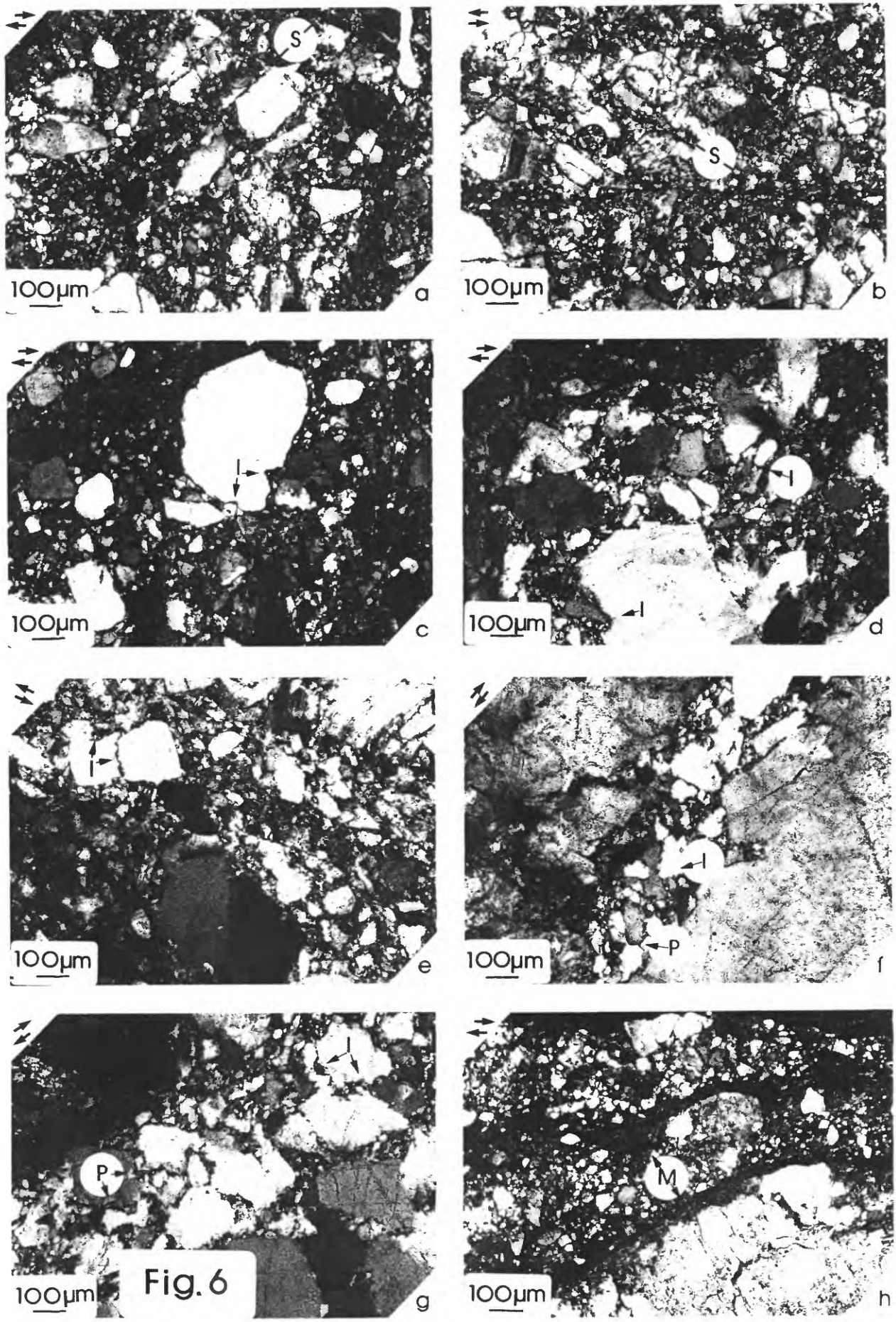

S
100µm
a
S
100µm
b
I
100µm
c
I
I
100µm
d
I
100µm
e
I
P
100µm
f
P
I
100µm
g
M
100µm
h

Fig. 6

shows that the two limits of the cracks are often indicated by two trails of fluid inclusions in the host mineral. Contrary to some veins in the Miocene shale (which show successive micro-opening and sealing), most of the fractures within the San Gabriel fault appear to have been opened and sealed only once. This perhaps indicates successive heterogeneous states of stress associated with possibly large earthquakes on this fault.

The sealing of interconnected cracks is probably sufficient to significantly reduce the permeability of the country rock. But there is even evidence for compaction within the cataclastic zones. The mechanism of this compaction is not easy to identify, however. Solution cleavage is found (S, Fig. 6a,b) within two conjugate fault gouges (C and C', Fig. 4 site 5). Equidirectional indentation of some grains by others (due to compaction) may also be seen in the main gouge, as well as in secondary gouges in the country rock (I, Fig. 6c,d,e). It should be noted that the mechanical indentation of grains has been experimentally reproduced (Deelman, 1976); this produces crushed material which remains in the vicinity of the indented zone. The field observations, in our case, do not show any evidence for this. Indentation by dissolution is then the most plausible explanation for these structures. In addition, the indentation of large grains (I, Fig. 6c) looks very similar to indented fossils created by pressure solution (Engelder, 1982). Finally, the interpenetration of some grains implies a change in shape of these grains. Evidence of undulated extinction revealing local plasticity is observed (P, Fig. 6f,g), but the change in shape is not always associated with such structures. Most of this change is associated with pressure solution (I, Fig. 6f-g). Partitioning of the mechanisms of deformation has thus been observed, where part of the compaction is attributable to pressure solution.

Some special features seem to appear in decimeter-sized blocks collected within the San Gabriel fault gouge. Mica minerals coat the limits of the secondary fault gouges (M, Fig. 6h). The growth of these minerals is also observed in pressure shadow zones around large quartz grains (Fig. 6h). These features indicate that along some paths of mass transfer, hydrothermal or metamorphic reactions may have contributed to the reduction of permeability (see also Wintsh et al, this volume). Nonetheless, these reactions are limited for the most part to areas associated with large fracture systems that have not been completely sealed. These types of reactions have been related to large fluxes of fluid flow (Chester et al., 1993).

## Conclusions from observations of naturally sealed cracks

Fluid inclusions indicate that a fluid phase was present during the sealing of the cracks. Observations show that most of the cracks were sealed due to the precipitation of input material; these cracks could not have self-healed. Systematic associations between crack sealing and solution cleavage, and the complete sealing of each crack, indicates that pressure solution is one of the primary mechanisms for sealing interconnected cracks near active faults.

In close proximity to the faults, the location of successive crack openings may be variable. Successive openings can occur in the same vein (crack-seal mechanism; see Ramsay, 1980), where this leads to mineral fibers parallel to the direction of constant displacement (part of the veins near the Little Pine fault). Alternatively, they can occur as different cross-cutting veins, each one being opened and sealed only once in a zone with heterogeneous (seismic) displacement (part of the veins near the Little Pine fault and most of the veins near the San Gabriel fault).

The large size of the closed system and the large distances associated with mass transfer (near the Little Pine fault) indicate that mass transfer by infiltration was coupled with diffusion-limited pressure solution. In the country rock, the small gouges associated with secondary faults were compacted after the initial cataclastic event. At least part of this compaction was accommodated by pressure solution. Within the main gouges, hydrothermal and/or metamorphic reactions occurred, indicating the presence of large fluxes of fluid.

## EXPERIMENTAL DEFORMATION AND SEALING PROCESSES

### Compaction and deformation in flow-through experiments

Methods: Deformation experiments using halite aggregates (8.8% initial porosity) were performed in a flow-through system at room temperature (Chen et al., 1992). Saturated NaCl solutions, maintained at a constant in-flow fluid pressure, were used as the infiltrating fluid. In addition, one run used ethanol as the infiltrating solution. A constant axial stress ($\sigma_1$ = 2 to 6 MPa) and constant confining pressure ($\sigma_3$ = 1 MPa) were also imposed on the system (Fig. 7a). During the experiments, several parameters were monitored: displacement of the piston (Fig. 7b), mass of the fluid at the output (Fig. 7c), and in-situ porosity (using an acoustic wave velocity scanner; Fig. 7d). After the experiments, the porosity and the X-ray tomographic density were measured (Fig. 8a-8d); these were supplemented by optical microscope and SEM observations (Fig. 9a,b).

Results: The overall effect of the flow-through experiments under applied stress was to decrease the permeability; our results are similar to those obtained in other flow-through experiments on quartz aggregates (Elias & Hajash, 1992; Blanpied et al., 1992) and during the initial stages of experiments on jointed marble (Marone et al., 1988). In our experiments, changes in permeability were strongly dependent on the deviatoric stress level. When the deviatoric stress was zero ($\sigma_1 = \sigma_3$), finger-shaped dissolution zones occurred, which during the first day, lead to a dramatic increase in the sample permeability (Fig. 7c). With the deviatoric stress ($\sigma_1 - \sigma_3$) ranging from 1 to 5 MPa ($\sigma_3$ = 1 MPa), the permeability initially increased with the appearance of finger-shaped dissolution zones, but then subsequently and rapidly decreased after the first few days (Fig. 7c,d). Deformation was associated with the changes in the permeability. In all of the experiments, the measured strain rates always approached a stable limiting value as a function of time (see Fig. 7b). In the presence of ethanol, the recorded deformation was an order of magnitude less than observed for similar experiments using saturated NaCl solutions (Fig. 7b), and the permeability remains almost stable with only a very slow decrease (Fig. 7c).

During the first few days of the experiments using NaCl saturated solutions, the extent of dissolution was indicated by a decrease in the acoustic wave velocity (Fig. 7d); this was subsequently confirmed by SEM observations of free-face dissolution (Fig. 9a). The initial stages of the experimental runs involved the rapid dissolution of some minor minerals (carnallite, sylvite) which were not in equilibrium with the original NaCl saturated solutions. In addition to this, dissolution reactions were driven by the enhanced solubility of surface fines created by the crushing of the grains (surface curvature-solubility effect) and by the increased elastic and plastic surface strain energy associated with the imposed stress (free-face pressure solution, Fig. 9a).

After the first few days, the permeability always decreased over a period of several months of NaCl solution infiltration, up to the point where the samples became impermeable to further fluid flow (Fig. 7c,d). The impermeability was induced by the progressive sealing of the paths of fluid flow. This can be seen in the X-ray tomographic density cross sections of the samples after different time durations of fluid flow (Fig. 8a-d). In these figures, low-permeability zones are white colored areas that laterally spread out over time (Fig. 8a = initial, 8b = 4 days, 8c & d = 2 months). Just before the complete sealing of the sample, infiltration was localized along single finger-shaped paths of transport. Traces of these localized transport pathways (observed as zones of high porosity) remained on some cross sections (Fig. 8d) even after circulation had stopped. The remaining porosity (non-connected) may decrease over time, due to the action of slow mass transfer processes, such as diffusion. SEM observations of the samples (post-experimental) also show that the open paths of fluid flow (Fig. 9a) were side-by-side with zones of pressure solution

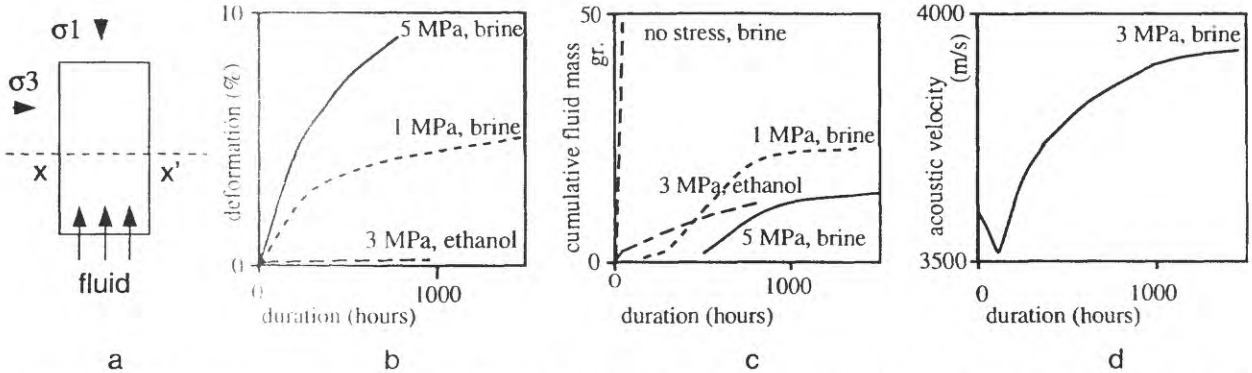



Figure 7: Experimental deformation of halite in a flow-through system: (a) sketch of the experimental system with stress and fluid flow directions (arrows). (b) deformation (c) cumulative mass fluid collected at output (d) acoustic wave velocity versus time. Stress values are the deviatoric stress ($\sigma_1 - \sigma_3$). Brine is the NaCl saturated solution.

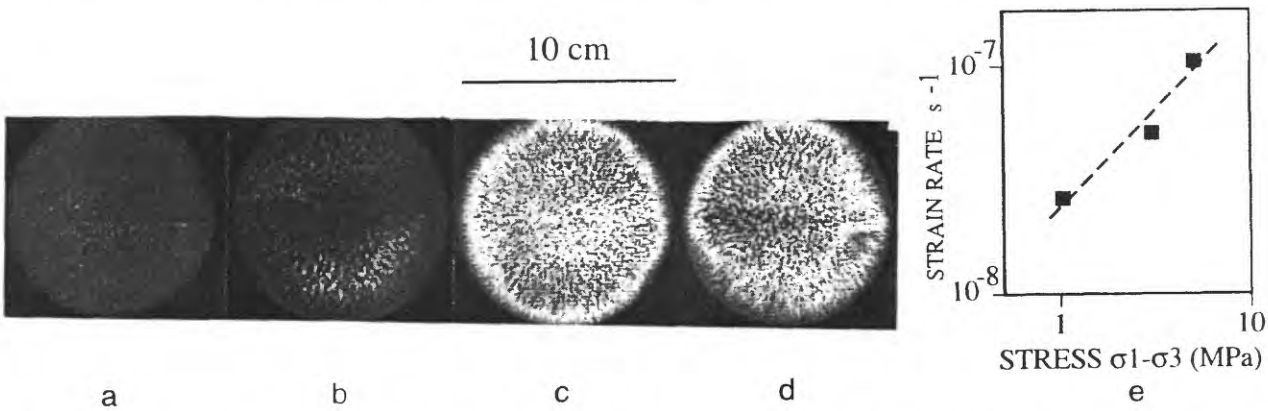



Figure 8: X-ray tomographic density image of the sample (white = low porosity, black = high porosity), location of the cross sections (a), (b) & (c) in Fig. 7a (XX'): (a) initial (b) after 4 days (c) & (d) after 2 months of fluid flow (e) Strain-rate versus stress relation for compacted aggregate (deformation with constant strain rate when most of the porosity is destroyed, see Fig. 7b).

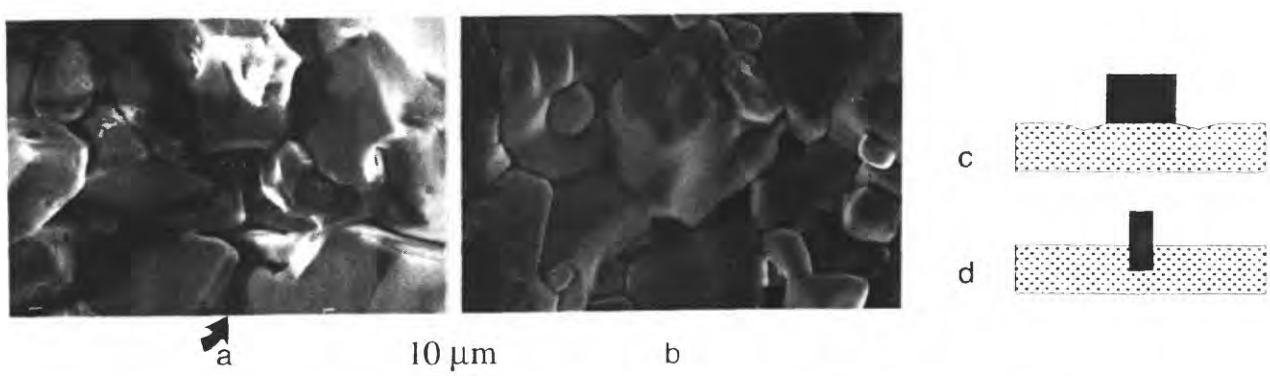



Figure 9: Scanning electron microscope (SEM) observations after fluid flow: (a) fluid transport path enhanced by free-face dissolution (arrow) (b) compaction of the grains by pressure solution (with associated water-film diffusion). Indenter experiments on halite: (c) free-face pressure solution, or (d) water-film diffusion pressure solution; the mechanism depended in part on the initial saturation state of the solution and on the diameter of the indenter.

compaction (Fig. 9b). Observations show that dissolution occurred on all the faces under stress (whatever their orientation versus stress).

Discussion: The thermodynamic driving forces associated with compaction and deformation by pressure solution are a function of solution undersaturation, strain energy, and normal stress. The order of these factors is closely related to their chronological importance in the experiments, as is discussed below.

1 - Solution undersaturation: The initial contact between the NaCl-saturated brine and the halite salt core produced the rapid dissolution of minor secondary phases, such as sylvite (KCl) and carnallite ($KMgCl_3 \cdot 6H_2O$). The driving force for the dissolution reactions of these minerals is related to the degree of solution undersaturation. Thus, the Gibbs free energy ($\Delta G_{diss}$) driving the dissolution reaction (i.e. $AB \rightarrow A^+ + B^-$) for a given salt phase can be expressed as

$$\Delta G_{diss} = RT \ln \frac{a_A a_B}{K} \qquad (1)$$

where $a_A a_B$ is the activity product of aqueous species $A^+$ and $B^-$ at any given time, and K is the equilibrium constant for the dissolution reaction. The reaction is driven in the forward direction until the activity product equals the equilibrium constant. Based on the chemical analyses of the output solutions, the concentrations of $K^+$ and $Mg^+$ in solution exceeded those of $Na^+$ (by up to a factor of 4), indicating that the initial stages of dissolution favored the dissolution of sylvite and carnallite. This is due to the fact that the initial solutions were undersaturated with respect to these phases.

2 - Strain energy: The second phase of dissolution and deformation, which occurs on a time scale of days to months, is related to the plastic and elastic deformation of surfaces *surrounding* grain-to-grain contacts. The associated dissolution occurs on grain surfaces in contact with a free liquid . The strain energy associated with plastic and elastic deformation (Reuschlé et al., 1988) serves to increase the Gibbs free energy for the dissolution reaction in the following manner:

$$\Delta G_{diss} = (\Delta G^0_{diss} - \Delta E) + RT \ln K' \qquad (2)$$

where $\Delta G^0_{diss}$ is the Gibbs free energy for the reaction with all components in the standard state, $\Delta E$ is the strain energy associated with deformation, and K' is the solubility product of the strained phase. The Gibbs free energy of reaction, based on K', becomes

$$\Delta G_{diss} = RT \ln \frac{a_A a_B}{K'} \qquad (3)$$

Thus, the above reaction will proceed in the forward direction so long as the activity product $a_A a_B$ is less than the solubility product (K') of the strained phase. It is important to note that since $K' > K$, dissolution (which is equivalent to free-face pressure solution) can occur in saturated solutions only with respect to those grain areas that have been subjected to strain.

3 - Normal stress: Finally, dissolution and deformation by pressure solution reactions *between* grain contacts occurs very slowly, on the order of months to years. The effect of a normal stress on a solid serves to increase the driving force for dissolution by a quantity equal to the product of the normal stress on the solid ($\sigma_n$) and the molar volume ($V_\alpha$) (see Robin, 1978). Therefore, the Gibbs free energy driving pressure solution (i.e. dissolution *between* grain-to-grain contacts) can be expressed as

$$\Delta G_{diss} = (\Delta G^0_{diss} - \sigma_n V_\alpha) + RT \ln K'' \qquad (4)$$

where K" is the solubility product associated with the normal stress $\sigma_n$. The above relationship predicts an increase in the kinetics of dissolution (i.e. strain rate) at grain-to-grain contacts with an increase in the normal stress. However, as discussed earlier, the limiting reaction mechanism may not be a function of the rate of detachment of species at the surface/fluid interface, but rather may be limited by the rate of diffusion of product species in the trapped fluid film between the grain contacts (water-film diffusion). The value of the diffusion coefficient estimated from the aggregate deformation (see below, comments of Fig. 8e) may be compared with other standard values (see discussion). Precipitation of products occurs on surfaces in contact with a free liquid, or on those faces subjected to the minimum compressive stress.

The observed equidirectional dissolution of mineral grains, which seems to be a characteristic of the compaction process, has also been observed with other experiments (Spiers & Schutjens, 1990; Cox & Paterson, 1991), as well as on some natural structures (Tigert & Al Shaieb, 1990). This type of dissolution occurs as long as the fluid pressure remains less than $\sigma_3$. This is due to the fact that the chemical potential of faces perpendicular to both $\sigma_1$ and $\sigma_3$ is higher than the chemical potential of free faces (faces submitted to fluid pressure). The higher the chemical potential, the greater is the driving force for dissolution; this is due to the fact that the gradient in chemical potential drives pressure solution.

When the connectivity between the pores is destroyed, the fluid pressure in isolated voids increases to a value approaching $\sigma_3$. Mass transfer of material takes place between faces perpendicular to $\sigma_1$ (zones of dissolution) and faces perpendicular to $\sigma_3$ (zones of precipitation). This progressively leads to a change in grain shape (i.e deformation- see sketch of this evolution in Fig. 10a). With levels of finite compaction/deformation less than 9%, these changes in grain shape were not easily observed. Nonetheless, when considering the derived stress versus strain rate relation (Fig. 8e) obtained just before the complete sealing of the sample (last part of the strain versus time curve, with constant strain rate in Fig. 7b), this relation is linear and characteristic of pressure solution creep. At this latter stage, even if some isolated finger-shaped transport zones remained open (Fig. 8d), the majority of the sample had already undergone a significant degree of compaction (mean final porosity = 2.7%).

Finally, it is important to recall that no significant compaction was obtained on halite aggregates in the presence of ethanol. The relatively insoluble nature of halite in ethanol precluded its dissolution by pressure solution, either by free-face or by water-film diffusion. This is a confirmation that the compaction and deformation recorded in our experiments with NaCl saturated solutions was due to pressure solution, and did not occur via plastic deformation.

## Driving force and kinetics of mass transfer: dissolution under indenters

An indenter technique was used to perform pressure solution experiments on halite (Gratier, 1993) and quartz. The experiments permitted the differentiation of two different mechanisms of pressure solution-deposition, depending on the conditions of deformation. *Free-face pressure solution* (Fig. 9c) occurred rapidly (time scale of days to weeks) and was driven by elastic and plastic strain energy (Tada & Siever, 1986), as indicated by dissolution around the indenters (*i.e.* conical-shaped grooves, with internal diameter equal to the indenter diameter). This effect was only observed when the initial solution was slightly undersaturated. It was independent of the diameter of the indenters and it was probably reaction rate limited. Dissolution at the indenter/mineral interface, associated with *water-film diffusion* (Fig. 9d), occurred slowly (over months, to one year) and was driven by the difference in normal stress, as indicated by dissolution under the indenters (*i.e.*, cylindrical-shaped holes with diameters equal to the indenter diameter). The displacement rate was inversely proportional to the diameter of the indenter and was probably water-film diffusion rate limited. In this case, a linear relation was observed between the displacement rate and the deviatoric stress values (Gratier 1993); this result was analogous to the deformation relation observed for the halite aggregates (Fig. 8e).

The conclusion that can be drawn from such experiments is that the saturation state of the solution may change the relative importance of the various driving forces of mineral/water mass transfer. Since large and intermittent fluid flow can occur during rock deformation (Sibson, 1977; Etheridge et al., 1984; Kerrich, 1986), the rate-limiting mechanism associated with pressure solution may change from free-face reactions (when the solution is slightly undersaturated) to water-film diffusion (a process which continuously occurs, but at much slower rate).

### Conclusions from the experimental approach

Fig. 10a summarizes the proposed interpretation of how pressure solution mediated dissolution of halite aggregates occurs. In general, these conclusions are also valid for the experiments carried out using the indenter method.

1 - Free-face fluid/solid reactions (often reaction rate limited) usually occur rapidly, on the order of days to weeks. These reactions are driven by solution undersaturation with respect to minor secondary phases, by surface energy (dissolution of crushed grains), and by elastic and plastic strain energy. These reactions always lead to the dissolution of grains, and depending on the degree of solution supersaturation, will also lead to precipitation reactions. Nonetheless, only special conditions generally lead to systematic crack sealing at this initial step in the deformation process.

2 - Compaction by pressure solution (water-film diffusion-limited) occurs slowly (months to years), with dissolution occurring at grain-to-grain contacts (equidirectional dissolution) and precipitation occurring on grain surfaces in contact with a free liquid. This process occurs as long as the fluid pressure remains lower than the minimum compressive stress.

3 - Deformation by pressure solution (also water-film diffusion-limited) is developed with dissolution occurring at grain-to-grain contacts perpendicular to the maximum compressive stress and precipitation occurring at surfaces submitted to the minimum compressive stress. This process occurs also very slowly (months, see discussion below) after the pore pressure reaches the minimum compressive stress (due to pore sealing).

## DISCUSSION

### Successive crack seal mechanisms in the interseismic period

When integrating both natural and experimental observations, the following possible successive crack sealing processes may be expected to occur in active faults during interseismic periods (see sketch, Fig. 10b). Initially, an earthquake will serve to increase the overall permeability and reduce fluid pressures to levels approaching hydrostatic values within the fault zone.

The first stages of fluid/rock interactions can be characterized by rapid kinetics which are dependent on the reactions taking place on free faces (Fig. 10b-1). The nature of the crack aperture also plays an important role. Small aperture cracks can be self-healed, where the driving force for the reaction is the minimization of crack surface free energy (Smith & Evans, 1984; Brantley et al., 1990). Fluid/mineral interactions in larger aperture cracks can be driven by a variety of processes: rapid dissolution of extremely small grains created in cataclastic zones (Elias & Hajash, 1992), dissolution of surfaces with elevated elastic and plastic strain energy (Tada & Siever, 1986) dissolution of mineral phases not in equilibrium with the infiltrating solution (Chester et al., 1993). Systematic sealing by free-face reactions is only expected under special conditions, such as: cracks with a small aperture, the presence of supersaturated fluids, hydrothermal or metamorphic reactions which produce new minerals (Wintsh et al, this volume). The kinetics of free-face reactions are generally reaction rate limited in the case of low solubility minerals (silicates, oxides), which have high activation energies (50-90 kJ/mole/K: Lasaga, 1984;

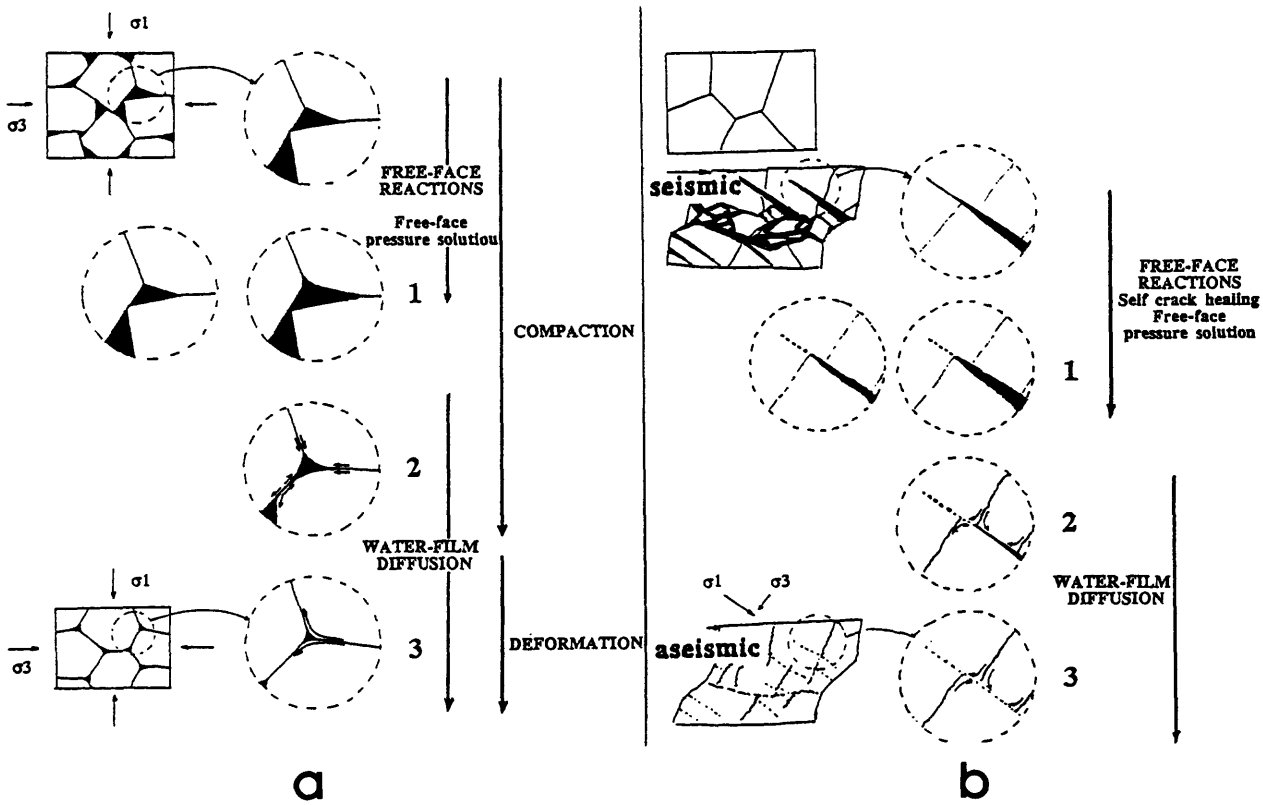



Figure 10: Possible evolution of the mechanisms of dissolution and precipitation (a) in aggregate and (b) in fractured rocks, solid is white, fluid is black. At the beginning of the compaction (a-1) or after an earthquake (b-1) free-face reactions occur rapidly due to effect of increased surface, plastic, and elastic energies (days to months). But only special conditions can lead to the sealing of the sample with such a mechanism. When pressure solution is driven by differences in normal stress along grain boundaries (a2-3) or around the blocks (b2-3), the rate of sealing will be determined by water-film diffusion with very slow kinetics (years to thousands of years).

Smith & Evans, 1984; Gratier & Jenatton, 1984; Gratz et al., 1990). The dissolution of high solubility minerals, however, may be limited by diffusion at the (free) fluid/mineral interface.

Pressure solution creep (Fig. 10b, 2 -> 3) is the process responsible for the progressive and subsequent complete sealing of cracks and fissures associated with fault zones. The gouges are compacted in part by pressure solution (see Fig. 10a-2 with related discussion). Dissolution of material at grain-to-grain contacts normal to the maximum effective stress, followed by water-film diffusion, eventually leads to the precipitation of dissolved material, where precipitation is initially favored on the free faces of grains (sealing and compaction), then on faces normal to the minimum compressive stress (deformation). Such diffusive mass transfer generally occurs in small closed systems (100 μm to mm scale, Fig. 1d). However, large-scale infiltration coupled with diffusion-limited pressure solution occurs on larger scales (Fig. 1c"), which is the reason why larger closed systems are amenable to sealing, as well. The progressive sealing of a system implies that fluid pressures should increase from hydrostatic to lithostatic levels with time.

## Kinetics of crack sealing by pressure solution

The estimation of the kinetics of crack sealing is possible using certain theoretical relations derived for pressure solution. A simplified geometry for crack sealing is given in Fig. 10b (step 2 -> 3). Dissolution occurs on surfaces normal to the maximum compressive stress, whereas precipitation occurs on the free faces parallel to this compressive stress (crack limits). A method for estimating the incremental displacement rate ($\dot{\gamma}$) for the dissolution of a stressed face is based on the following expression (Rutter, 1976):

$$\dot{\gamma} = \alpha D w c \Delta\sigma v / R T d^2 \qquad (5)$$

where D is the water-film diffusion coefficient for the rate limiting species, w is the width of the transport path, c is the solubility, v is the molar volume of the soluble species, $\Delta\sigma$ is the deviatoric stress (i.e. the difference in stress between the zone of dissolution and the zone of precipitation), R is the universal gas constant, T is the temperature (K), d is the diameter of the dissolution zone, $\alpha$ is a numerical constant of integration dependent on the geometry of the block (32 for a disk-shaped dissolution surface, Rutter, 1976).

The major unknown parameters for pressure solution are D and w. These values may be estimated from the quartz and halite experiments. Using Eqn. 5, Dw may be estimated from the indenter experiments on halite. A mean value of $4.4 \times 10^{-19}$ $m^3/s$ was determined. From the flow-through halite aggregate experiments, a value of $6.6 \times 10^{-20}$ $m^3/s$ was obtained. These values may be compared with those obtained by Hickman (1989): Dw = $10^{-19}$ $m^3/s$ for the dissolution/deformation of halite at the same range of stress. The deformation of quartz aggregates (at 350°C in a 1.0 N $NaOH/H_2O$ solution) yielded perhaps more reliable values for crustal deformation. The values obtained ranged from $2.5 \times 10^{-19}$ to $2.5 \times 10^{-20}$ $m^3/s$ (Gratier & Guiguet, 1986). Experiments using indenter techniques (unpublished data on quartz) gave values of about $10^{-19}$ $m^3/s$.

Dw values ranging from $10^{-19}$ to $10^{-20}$ $m^3/s$ seem to be a rough approximation for quartz at depths corresponding to temperatures of 350°C. The other parameters necessary for applying Eqn. 5 to natural deformation may be taken from the literature. The solubility of quartz in solution at 350°C can be derived from log K = -1.89 (Rimstidt and Barnes, 1980). Based on the geometry of naturally deformed rocks around faults, the dimensions associated with the dissolution zone may be estimated. It is commonly observed that in natural systems mass transfer is not limited to the scale of single grains, but rather is constrained to a multigrain scale. This scale corresponds to multigranular blocks that are limited in size by solution cleavage (at the base) and cracks (sides)

(Gratier, 1987; Gratier & Gamond, 1990). A commonly observed crack width is on the order of 10 μm (see Fig. 5). The mean diameter of the zone of dissolution (d) around each crack can be estimated to range from 100-500 μm. Using an assumed deviatoric stress value of 50 MPa, the time needed for complete crack sealing ranges from 5.6 to 56 years (with d = 100 μm), and 140 to 1400 years (for d = 500 μm). Of course, a more accurate estimation depends on knowledge of both the exact geometry of the dissolution and deposition zones and on the use of more accurate values for the diffusion coefficient (D) and the width of the water-film (w).

In conclusion, the kinetics of crack sealing by pressure solution can be quite variable. The rate of crack sealing may be quite rapid if dissolution occurs on free faces (Fig. 1c', reaction rate-controlled). However, when kinetically-limited by diffusion through a trapped water-film, the rate of diffusion may constrain the rate of increase of fluid pressure within faults. In such a case (Fig. 1c", 1d), pressure solution is not a rapid process relative to other geodynamic processes which can lead to elevated fluid pressures. The estimated values for pressure solution crack sealing are of the same order of magnitude as the recurrence time of earthquakes.

### Strain softening and hardening associated with pressure solution

Since pressure solution on polymineralic aggregates can induce chemical differentiation, this process is associated with strain softening and hardening. The strength evolution of rocks can be considered both in terms of time and space.

Spatial evolution. Most soluble minerals (quartz, feldspar, calcite) are mechanically stronger than insoluble micas and clays (minerals which are typically alteration products). The selective removal of soluble minerals decreases the strength of rocks and leads to a progressive localization of dissolution along cleavage seams (Cosgrove, 1976; Robin, 1979). On the contrary, precipitation of minerals increases the strength of rocks and leads to zones of mineralization that are stable with respect to further dissolution . This mineralogical differentiation occurs in active faults between the gouge zone (zone of dissolution with decreased quartz and feldspar contents) and the country rock (zone of crack sealing with increased quartz and feldspar contents).

Time evolution: The dependence on time is a function of the geometry of the features associated with pressure solution. When pressure occurs with parallel solution cleavage normal to planar veins (Fig. 3, 4, 5, 6a-b) the geometry is stable and the banding (layer with dissolution / layer with precipitation) evolves with time up to an equilibrium state which depends on the conditions of deformation (Gratier, 1987). On the contrary, when dissolution/precipitation occurs at the scale of a grain (Fig. 6c-g), the time evolution depends on the geometry and on the type of pressure solution mechanism. For the case of a monomineralic aggregate under a constant mean deviatoric stress, free-face dissolution reduces the overall grain size (Dewers & Ortoleva, 1990). The decrease in the grain-to-grain surface area (Fig. 9a) increases the local stress imposed on the areas of contact and leads to an accelerated dissolution rate; this can be considered to lead to strain softening. Counteracting this effect is the process of progressive grain-to-grain indentation. The associated increase of grain-to-grain surface decreases the local stress on the contacts and leads to progressive strain hardening of the aggregate (Fig. 6 and 9b).

## CONCLUSIONS

Experiments and observations of natural structures show that various types of solid/fluid reactions can occur during interseismic periods. The kinetics of these reactions depends on the nature of the limiting process. Solid/fluid reactions may be divided into two categories:

Reactions which occur between a solid under stress and a free fluid phase. These types of reactions include self-crack healing, dissolution of crushed grains, free-face pressure solution (driven by strain energy), metasomatic reactions. Most of these processes are reaction rate limited, where the slow step in the process is related to detachment or precipitation kinetics. Of course, highly soluble minerals may dissolve in a diffusionally-limited manner. Time scales for these free-face reactions are on the order of days to months. External conditions associated with deformation and mineral/fluid interactions (temperature, pressure, fluid and solid compositions, fluid flow rates) determine whether dissolution occurs by itself, or whether it is subsequently followed by localized precipitation.

Observations show that in active faults most of the cracks are sealed due to the input of material and have not self-healed. Hydrothermal and metamorphic reactions have been observed to occur within gouge zones, but they do not seem to be associated with the sealing of transport paths. Only special conditions can therefore produce systematic and rapid crack sealing with free-face reactions.

Reactions which occur within trapped fluid phases between two solids under stress. These reactions include compaction and crack sealing (driven by the difference between the normal stress at grain-to-grain contacts and the surfaces of grains in contact with free fluids present in interconnected voids and cracks). The dissolution process is postulated to be limited by the rate of diffusion in the trapped water-film. These rates are typically much slower, with crack sealing occurring on the order of years to thousands of years. In this case mass transfer can be either water-film diffusion (Fig. 1d), or fluid infiltration coupled with water-film diffusion (Fig. 1c").

The systematic association between crack sealing and solution cleavage indicates that pressure solution-deposition reactions are the primary mechanism for sealing interconnected cracks near active faults. This type of crack sealing, even though it occurs on a micrograin scale, can completely seal a fracture system at a much larger scale, due mainly to widespread fluid circulation and infiltration. The compaction of gouges after the original cataclastic event can in part be attributed to pressure solution.

Some consequences of pressure solution crack sealing. If pressure solution is rate limited by water-film diffusion, then the kinetics of crack sealing are slow enough to be an important process in controlling the increase in fluid pressure around faults during interseismic periods. It also has an important bearing on the associated mechanisms of critical failure. The time evolution of pressure solution, however, not only depends on the operative mechanisms, but also on geometrical constraints which determine the scale over which the process is effective (solution cleavage scale vs. single grain dissolution).

Deformation associated with pressure solution very often induces the chemical and mechanical differentiation of rocks. Gouge zones where preferential dissolution occurs generally lose mechanical strength in comparison to zones (country rock) where associated precipitation reactions take places. The progressive decrease in fluid flow through faults occurs in parallel with the progressive decrease of the post-seismic deformation rate due to pressure solution, perhaps to a near constant creeping rate (with mass transfer from cleavage to vein) when the fault fracture system has been completely sealed. This fits in well with the overall deformational behavior of the upper crust, where brittle behavior (cataclastic faults, seismic events, fast strain rates) alternates with ductile behavior (solution cleavage, aseismic creep, slow strain rates).

**Acknowledgements**: The experiments were carried out in Grenoble. Funding was provided by CNRS-INSU programs and ANDRA. The studies in California were supported by the Southern California Earthquake Center. We thank C. Nicholson for a helpful review and B. Patrick, R. Guiguet, and L. Jenatton for their technical assistance. Institute for Crustal Studies Contribution no 152-37TC.

## References

Arthaud F. & M. Mattauer, 1969: Exemples de stylolites d'origine tectonique dans le Languedoc, leur relation avec la tectonique cassante. *Bull. Soc. Geol. Fr. 11,* 738-744.

Bernard D., J.P. Gratier & A. Pêcher, 1977: Application de la microthermométrie des inclusions fluides des cristaux syncinématiques à un problème tectonique. *C.R. Somm. Soc. Geol. Fr. 5,* 284-288.

Berry F.A.F., 1973: High fluid potentials in the California Coast Range and their tectonic significance. *Am. Ass. Petrol. Geol. Bull. 57,* 1219-1249.

Boullier A.M. & F. Robert, 1992: Paleoseismic events recorded in Archaen gold-quartz vein networks, Val d'Or, Abitibi, Quebec, Canada. *J. Struct. Geol. 14,* 161-179.

Blanpied M.L, D.A. Lockner & J.D. Byerlee, 1992: An earthquake mechanism based on rapid sealing of faults. *Nature 358,* 574-579.

Brace, W.J., 1972: Pore pressure in geophysics. In: *Flow and fracture of rocks, The Griggs Vol.* Am. Geophys. Union.

Brantley S.L., B. Evans, S. H. Hickman & D.A. Crerar, 1990: Healing of microcracks in quartz: implication for fluid flow. *Geology 18,* 136-139.

Brimhall G.H. & D.A. Crerar, 1987: Ore fluids: magmatic to supergene. In: thermodynamic modeling of geological materials, Carmichael & Eugster eds. *Reviews in Mineralogy, 17,* 235-254.

Byerlee J.D., 1990: Friction, overpressure and fault normal compression. *Geophys. Res. Lett.,* 17, 2109-2112.

Cathles L.M., 1977: An analysis of the cooling of intrusives by ground water convection which includes boiling. *Econ. Geol., 72,* 804-826.

Chen T., J.P. Gratier & R. Hellmann, 1992: Evolution de la perméabilité de géomatériaux sous contrainte liées à des circulations de fluides (dissolution-colmatage). *ANDRA, internal report,* 81p.

Chester F.M., J.P. Evans & R.L. Biegel, 1993: Internal structures and weakening mechanisms of the San Andreas fault. *J. Geophys. Res.* 98, 771-786.

Cosgrove J.W., 1976: The formation of crenulation cleavage. *J. Geol. Soc. Lond. 132,* 155-178.

Cox S.F. & M.A. Etheridge, 1983: Crack-seal fibre growth mechanism and their significance in the development of oriented layer silicate microstructures. *Tectonophysics, 92,* 147-170.

Cox S.F. & M.S. Paterson, 1991: Experimental dissolution-precipitation creep in quartz aggregates at high temperatures. *Geophys. Res. Lett., 18,* 1401-1404

Deelman J.C. 1976: Lithification analysis, experimental observations. *Geol. Rundshau, 65,* 1055-1078.

Dewers T. & P. Ortoleva, 1990: A coupled reaction transport mechanical model for intergranular pressure solution and cementation of clean sandstones.*Geochim. Cosmochim. Acta, 54,* 1609-1625.

Elias P.E. & A.Hajash Jr., 1992: Changes in quartz solubility and porosity due to effective stress: an experimental investigation of pressure solution. *Geology, 20,* 451-454.

Engelder T. 1982: A natural example of the simultaneous operation of free-face dissolution and pressure solution. *Geochim. Cosmochim. Acta, 46,* 69-74.

Etheridge M.A., V.J. Wall, S.F. Cox & R.H. Vernon, 1984: High fluid pressure during regional metamorphism and deformation: implication for mass transfer and deformation mechanism. *J. Geophys. Res, 89,* 4344-4358.
Fyfe W.S., N.J. Price & A.B. Thomson, 1978: Fluids in the earth crust. Elsevier, Amsterdam.
Gratier J.P. & L. Jenatton, 1984: Deformation by solution-deposition and re-equilibration of fluid inclusions in crystals depending on temperature, internal pressure and stress. *J. Struct. Geol., 6,* 189-200.
Gratier J.P. & R. Guiguet, 1986: Experimental pressure solution on quartz grains: the crucial effect of the nature of the fluid. *J. Struct. Geol., 8,* 845-856.
Gratier J.P. 1987: Pressure solution creep and associated tectonic differentiation in sedimentary rocks. *Geol. Soc. Lond. Spec. Publ. 29,* 25-38.
Gratier J.P. & J.F. Gamond, 1990: Transition between seismic and aseismic deformation in the upper crust. In: Deformation mechanisms, rheology and tectonics, edited by Knipe & Rutter *J. Geol. Soc. Lond., Spec. Publ. 54,* 461-473.
Gratier J.P., 1993: experimental pressure solution of halite by an indenter technique.*Geophys. Res. Lett.* in press.
Gratz A.J., G.P. Bird & G.B. Quito, 1990: Quartz dissolution kinetics in aqueous basic solution: surface kinetics of "perfect" crystallographic faces, *Geochim. Cosmochim. Acta 54,* 2911-2922.
Hedinquist J.W., G. Reyes, S.F. Simmons & S. Taguchi, 1992: The thermal and geochemical structures of geothermal and epithermal systems: a framework for interpreteting fluid inclusion data. *Eur. J. Mineral. 45,* 989-1015.
Hickman S., 1989: Experimental studies of pressure solution and crack healing in halite and calcite. Ph.D. thesis, Mass. Inst. Tech., Mass.
Hickman S.H., 1991: Stress in the lithosphere and the strength of active faults. *Rev. Geophys., 29,* 759-775.
Hubbert M.K. & W.W. Rubey, 1959: Mechanics of fluid filled porous solid and its application to overthrust faulting. *Geol. Soc. Amer. Bull. 70,* 115-166.
Ildefonse J.P. 1980: Recherche sur le transport du silicium et du magnésium en conditions hydrothermales. *Thèse d'Etat,* Orléans.
Jennings C.W. 1977: Geological map of California, 1/750,000. D.M.G. California.
Kerrich R., 1986: Fluid infiltration into fault zones. *Pure Appl. Geophys. 124,* 225-268.
Kirby S.H. 1983: Rheology of the lithosphere. *Rev. Geophys. Spac. Phys. 21,* 1458-1487.
Lasaga A.C., 1984: Chemical kinetics of water-rock interactions. *J. Geophys. Res., 89,* 4009-4025.
Lachenbruch A., 1980: Frictional heating, fluid pressure and the resistance of fault motion. *J. Geophys. Res., 85,* 6097-6112.
Lefaucheux F. 1988: Croissance expérimentale de calcite dans un flux orienté, extended abstract, Conférence sur les interaction fluide/minéraux, Nancy.
Lloyd G.E. & R.J. Knipe, 1992: Deformation mechanisms accommodating faulting of quartzite under upper crustal conditions. *J. Struct. Geol., 14,* 127-143.
McCaig A.M., & R.J. Knipe, 1990: Mass-transport mechanisms in deforming rocks: recognition using microstructural and microchemical criteria. *Geology, 18,* 824-827.
Marone C., J. Rubenstone & T. Engelder, 1988: An experimental study of permeability and fluid chemistry in artificially jointed marble. *J. Geophys. Res., 93,* 13763-13775.
Marquer D. & M. Burkhard, 1992: Fluid circulation, progressive deformation and mass transfer processes in the upper crust: the example of basement-cover relationships in the External Crystalline Massifs, Switzerland. *J. Struct. Geol.,14,* 1047-1057.
Poty B., H.A. Stalder & A.M. Weisbrood, 1974: Fluid inclusions studies in quartz from fissures of western and Central Alps. *Scheiwz. Min. Petr. Milt. 54,* 717-752.
Ramsay J.G., 1980: The crack seal mechanism of rock deformation. *Nature, 284,* 135-139.

Raj R., 1982: Creep in polycrystalline aggregates by matter transfer through a liquid phase, *J. Geophys. Res., 87,* 4731-4739.
Reuschlé T., T. Trotignon & Y. Gueguen, 1988: Pore shape evolution by solution transfer: thermodynamics and mechanisms. *Geophys. J. 95,* 535.547.
Rice J.R. 1992: Fault stress state, pore pressure distribution, and the weakness of the San Andreas fault. In: Fault mechanics and transport properties in rocks, edited by B. Evans & T.F. Wong, 475-503, Acad. Press.
Rimstidt J.D. & H.L. Barnes, 1980: The kinetics of silica-water reactions.*Geochim. Cosmochim. Acta, 44*, 1683-1699.
Robin P.Y., 1978: Pressure solution at grain to grain contacts.*Geochim. Cosmochim. Acta, 42*, 1383-1389.
Robin P.Y., 1979: Theory of metamorphic segregation and related processes.*Geochim. Cosmochim. Acta, 43,* 1587-1600.
Rutter E.H., 1976: The kinetics of rocks deformation by pressure solution. *Phil. Trans. R. Soc. Lond. A283,* 203-219.
Rutter E.H., 1983: Pressure solution in nature, theory and experiment. *J. Geol. Soc. Lond., 140,* 725-740.
Shimamoto T., Y. Kanaori & K. Asai 1991: Cathodoluminescence observations on low-temperature mylonites: potential for detection of solution-precipitation microstructures. *J. Struct. Geol. 13*, 967-973.
Sibson R.H., 1977: Fluid flow accompanying faulting: field evidence and models. Earthquake prediction: *Maurice Ewing Series,* edited by D.W. Simpson & P.G. Richards, AGU, Washington, D.C., vol. 4, 593-603.
Sibson R.H., 1987: Earthquake rupturing as a mineralizing agent in hydrothermal systems. *Geology, 15,* 701-704.
Sibson R.H., 1992: Implication of fault valve behaviour for rupture nucleation and recurrence. *Tectonophysics, 211,* 283-294.
Sleep N.H, & M.L. Blanpied, 1992: Creep, compaction and the weak rheology of major faults. *Nature, 359,* 687-692.
Smith D.L. & B. Evans, 1984: Diffusional crack healing in quartz.*J. Geophys. Res., 89,* 4125-4135.
Spiers C.J. & P.M. Schutjens, Densification of crystalline aggregates by fluid phase diffusional creep. In: Deformation Processes in Minerals, edited by D.J. Barber & P.G. Meredith, Unwin Hyman, London, 334-353, 1990.
Sprunt E.S. & A. Nur 1977: Microcracking and healing in granite: new evidence from cathodoluminescence. *Science, 205*, 495-497.
Tada R. & R. Siever, Experimental knife-edge pressure solution of halite, *Geochim. Cosmochim. Acta, 50,* 29-36, 1986.
Tigert V. & Z. Al-Shaieb, 1990: Pressure seals: their diagenetic banding patterns. *Earth-Sciences reviews, 29,* 227-240.
Weyl P.K. 1959: Pressure solution creep and force of crystallization, a phenomenological theory. *J. Geophys. Res. 64,* 2001-2025.
Wilkins R.W.I., J.P. Gratier & L. Jenatton, 1985: Experimental observation of healing of cracks and the formation of secondary inclusions in halite and quartz. *Eur. Cur. Res. Fluid Inclusions,* Abs.

# Micromechanisms of deformation and fluid behaviour during faulting.

R. J. Knipe,
Dept. of Earth Sciences, The University, Leeds, LS2 9JT, U.K.

Although the behaviour of fault zones is known to depend upon the interaction of deformation processes and fluids (Langsath and Moore, 1990; Carter et al, 1991), the details of these interactions are still poorly defined. Future modelling of fluid behaviour in fault zones requires quantification of the mechanical and chemical effects fluids have on fault zone evolution and dynamics. The range of dilation patterns associated with different deformation mechanism paths possible in fault zones impact differently on the mechanical behaviour as well as on fluid flow processes (Etheridge et al, 1984; Sibson, 1992). Microstructural analysis of fault rocks can provide crucial information for the identification micromechanical processes which are influenced by and in turn influence fluid behaviour.

This short paper emphasises three aspects of interaction of micromechanical processes and fluid behaviour which are important to the understanding of fault zone dynamics. These are:

1) That competition and interaction between damage induced by crystal distortion and fluid pressure cycles in fault zones can produce complex time-dependant failure patterns which are not incorporated into present models of fault behaviour. An important consequence of this is that periods of low fluid pressure (not just high fluid pressure) become crucial to the mechanical behaviour of the aggregate.
2) That 'creep' voids may play a role in the initiation of mechanical instabilities which affect fluid redistribution processes and fault zone behaviour near the quasi-plastic to elastico-frictional transition.
3) That fault zone ageing and sealing processes control the dual role of faults as fluid pathways and fluid flow barriers.

Each of these aspects is discussed in separate sections below and examples based on microstructural analysis (TEM, SEM) of natural fault rocks, developed under a range of upper crustal conditions, are presented. The analysis has included assessment of the progressive development of; cataclasites from porous (<5-20%) sandstones and quartzites deformed at depths of less than 4km and at temperatures of below 250$^{o}$C, and the evolution of transient fracture systems in mylonites undergoing progressive deformation through the quasi-plastic/elastico-frictional transition in the temperature interval 350-250 $^{o}$C. The research is part of a programme aimed at identifying the processes involved in the development of high level faulting, the role of diffusive mass transfer and crystal plastic processes during the evolution of fault zones, and the interaction between deformation and fluid flow.

## 1) Competition and interaction between crystal plastic induced damage accumulation and fluid pressure cycles during high level faulting.

Analysis of progressive evolution of cataclasites in sandstones and quartzites has revealed that low temperature ductile fracture processes can be important in the development of

microfractures and in the continued comminution of cataclasites (Lloyd and Knipe, 1991). Models of high level faulting (at temperature below 300°C) have concentrated on the role of brittle fracture, while the role of crystal plastic processes have been considered unimportant.

TEM and SEM analysis of point contacts between detrital grains in sandstones and quartzites involved in the evolution of fault zones illustrates a marked increase in the dislocation density and the number of dislocation cells/sub-grains. These cold work damage features help focus the location of microfracturing via a ductile fracture mechanism (Lloyd and Knipe, 1991). TEM of the microstructure in very fine grained (1-5μm), quartz rich cataclasites at various stages of sealing or healing also reveals the cataclastic fragments have a higher dislocation density than the undeformed host rock. Probably the most important microstructure found in these low temperature cataclasites is a mantle of crystal damage in the fragments. Figure 1 shows an example from a cataclasite generated from a Permian sandstone at temperatures estimated to be below 200°C. The TEM images reveal a range of damage features which mantle the small cataclastic fragments. These features are important because:

1) The subsequent fracture behaviour of the aggregate will be affected by these mechanical skins to the fragments. The presence of such zones will introduce a time-dependent, damage/wear accumulation factor into the fracture behaviour of the aggregate. That is, failure and instability initiation will depend on the accumulation of damage to a critical level. The size, depth of penetration and intensity of the damage layer will be controlled by deformation conditions (strain magnitude, stress levels and temperature), together with the mineralogy and crystal lattice orientations across contacts between the fragments in the aggregate. The damage mantle provides a microstructure which by selective removal enhances wear and generation of ultra-fine-grained fault products.
2) The damage zone will also be more susceptible to dissolution and reaction between or after high slip-rate events in the fault zone and will therefore have a pronounced influence on the sealing or ageing of the fault rock by diffusive mass transfer (DMT) induced compaction. The extent of the damage mantle will influence the volume fraction of the fragments which are susceptible to enhanced DMT and a decreasing rate of creep is likely as the wear mantles are removed after each slip event. Such a healing process has a crucial affect on both the strength and permeability properties of the fault zone. This susceptibility to dissolution also means that the damage microstructures are not stable features in many fault zones but are likely to be removed by the late stage straining.

These studies have confirmed that dislocation activity can be involved in damage accumulation in sandstones and in the wear damage induced by repeated deformation associated with the development of cataclasites. Future research needs to assess the controls on the evolution and stability of the damage/wear microstructures.

There is also the potential for an important interdependence between the damage accumulation and the fluid pressure cycles which can induce failure in gouge zones generated by cataclasis (Figure 2). The rate of damage accumulation at grain contacts will depend on the effective stress cycles associated with fluid pressure fluctuations in the fault zone. Stresses will be focused onto grain contacts during the ***high*** effective stress periods; i.e. during periods of low fluid pressure. Repeated fluid pressure cycles, below the levels needed to induce hydrofracture, may therefore induce damage accumulation and eventual rupture.

This model suggests that failure in the system need not be at peak fluid pressure and that the low fluid pressure time periods are not 'rest' periods in the mechanical behaviour of the system but time periods when important damage accumulation and thus changes to the mechanical behaviour can be induced.

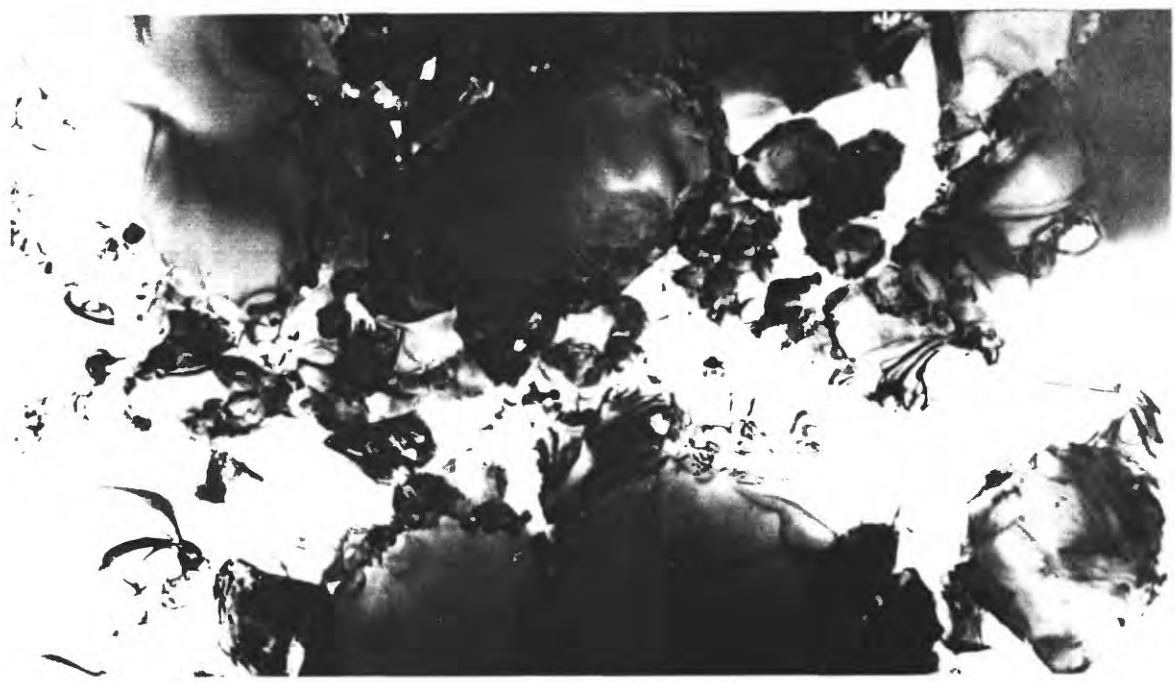


Figure 1. TEM micrograph of a natural fine grained cataclasite at an early stage of development. Note the mantle of damage features on the fragments. The distortion is associated with either increased lattice bending or concentrations of dislocations. Scale: 1.5 cm = 1μm.

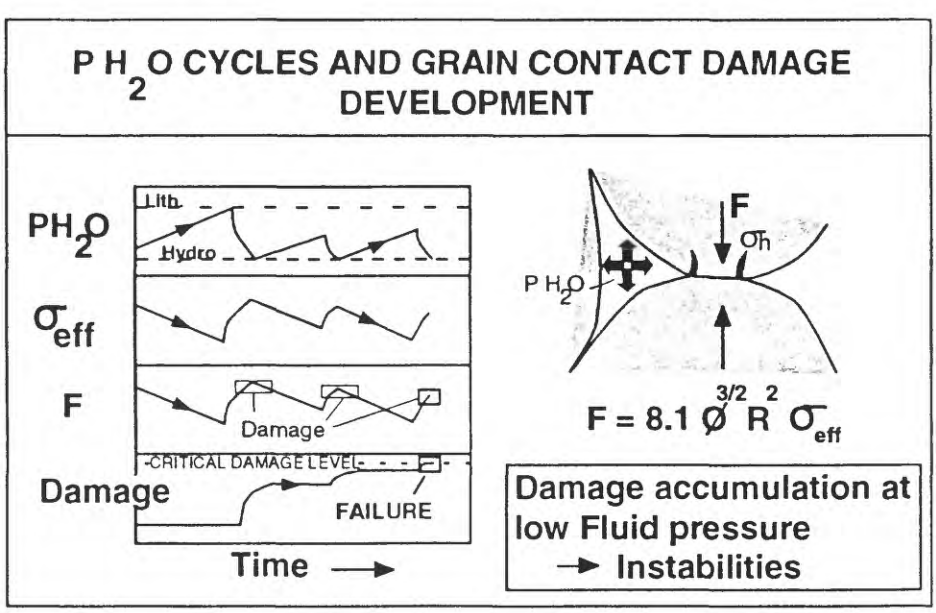



Figure 2. Model for the interaction of fluid pressure cycles, damage accumulation and failure in fault zones. The rate of damage accumulation at grain contacts will be dependant on the effective stress cycles associated with fluid pressure fluctuations in the fault zone. Stresses will be focused onto grain contacts during the high effective stress periods. Repeated fluid pressure cycles, may therefore induce damage accumulation and eventual rupture.

**2) Role of creep voids in fluid flow in fault zones active near the quasi-plastic to elastico-frictional transition.**

The transition from deformation dominated by crystal plastic creep processes (e.g. at high temperatures/slow strain rates), to lower temperature/high strain rate deformation, dominated by fracture and frictional sliding processes is recognised as an important breeding ground for earthquakes (Sibson, 1989). However, the detailed interaction of deformation processes in this quasi-plastic to elastico-frictional (Sibson, 1986) transition zone, which can lead to instability development, fracturing and fluid ingress are not well characterised. The gradual decrease in temperature which accompanies uplift and deformation in mylonite zones associated with large displacement contractional faults is an ideal location for studying this transition.

The transition to fracture dominated deformation from dislocation creep is preserved in the quartz mylonites from the Moine Thrust Zone of N. W. Scotland, as a result of the change in fault activity from the Moine Thrust to underlying, later thrusts (Knipe, 1991). The late stage evolution of these mylonites involves the repeated development of minor fracture arrays (quartz veins), some of which are then deformed by dislocation creep or affected by diffusive mass transfer processes. This transition of deformation processes is also accompanied by a change in the deformation kinematics, from one dominated by shear in the thrusting direction to one where there is a complex mixture of; i) backthrusting, ii) normal faulting back down the thrust plane and iii) vertical shortening accompanied by extensional veining as well as iv) continued localised shearing in the main thrust direction. The sequence appears to represents the death (the end of displacement activity) of the Moine Thrust under conditions where the transition to fracture dominated deformation has just commenced.

TEM /SEM of these mylonites has revealed that selected grain boundary orientations are characterised by a concentration of "creep voids" (Figure 3). These inclusions are up to 2μm in diameter, occasionally contain chlorite crystals (indicating fluid ingress into the pure quartzite), and are aligned along grain boundaries which are at high angles to the mylonitic fabric. These boundaries are those which are expected to experience extension during the late stage deformation sequence described above.

The growth and linking of such 'voids' is the probable cause of fracture initiation in the mylonites during deformation under retrogressive conditions and provide a fracture mechanism for the ingress of fluid. The void growth is related to the strain incompatibilities arising between grains, which have different lattice orientations and thus different degrees of deformability by dislocation processes. The decreasing temperature experienced by the aggregate reduces dislocation activity and effectively progressively removes the strain accommodation possible by crystal plastic processes. An imposed transient increase of the strain-rate in an aggregate deforming by dislocation creep would also induce the development of the creep voids.

This mode of fracture initiation is important to instability nucleation and the creation of fluid migration pathways (Figure 4).

The amount of dilation and the pattern of dilation changes induced by 'creep void' processes and the affect on fluid flow will depend upon the amount of deviation from the conditions where 'steady state' deformation by dislocation creep is possible or where strain accommodation by diffusive mass transfer prevents the creation of extended and linked micro-dilation sites. The density, migration velocity, distribution and linking probability of the voids, together with the competition between growth and closure by diffusive mass transfer, will dictate the influence on the mechanical behaviour. Where deformation conditions induce only a small deviation away from the situation where 'void' formation and grain boundary dilation is suppressed, small, isolated but migrating grain boundary voids and tubes may be created which

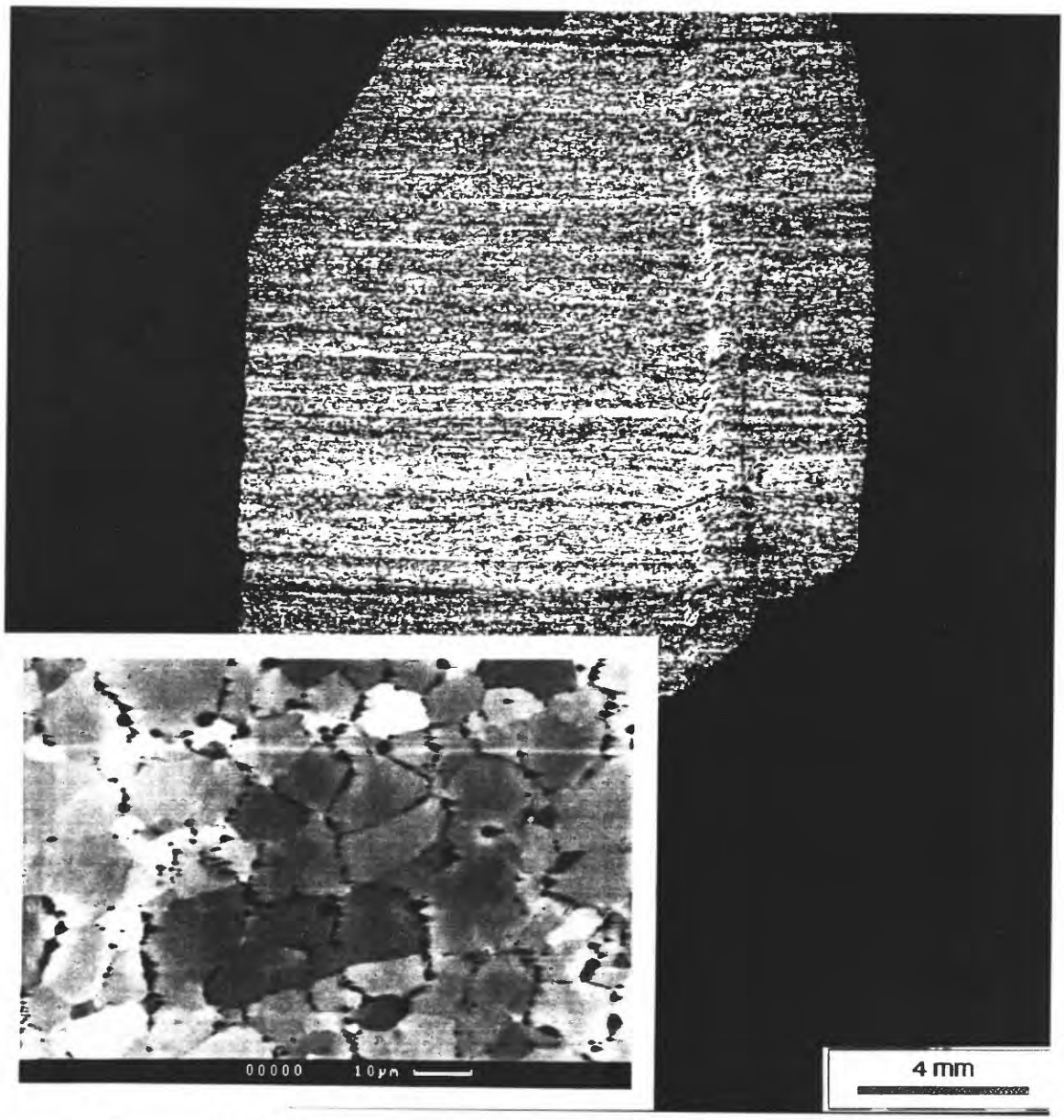



Figure 3. Micrographs illustrating the microstructure of a mylonite containing a late stage quartz vein. The inset is an backscatter SEM micrograph illustrating the concentration of creep 'voids' on extensional grain boundaries which lead the nucleation of the fractures and veins present.

link only occasionally. Larger deviations in the deformation conditions away from those which suppress 'void' development will lead to the creation of more extensive arrays of inclusions which are linked for longer periods. These conditions may allow fluid flow to be associated with pervasive flow through the deforming aggregate, although not all the boundaries need to be open at any one time. That is, while fluid-rock interactions may affect all grains, the active fluid migration pathway through the aggregate is not constant, but changes as previously active routes are closed or cemented up. This dynamic creation and destruction of fluid migration pathways,

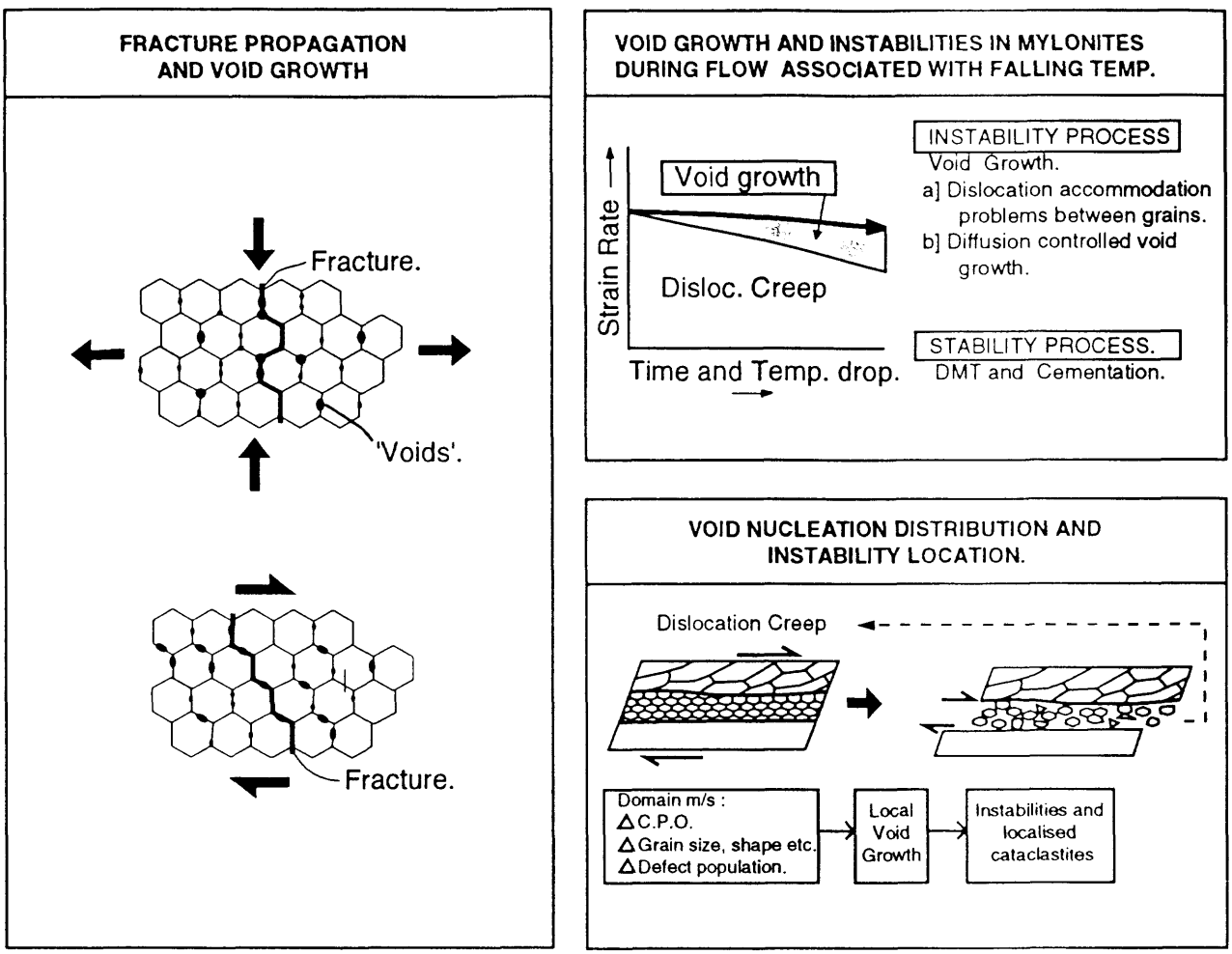



Figure 4. Model for the development of fractures and instabilities arising from the growth of creep 'voids' in a deforming aggregate.

where only a small percentage of the grain boundaries act as pathways, is important as it allows the aggregate to maintain it's strength. This dynamic situation will have a specific range of conditions (strain-rates, stress levels, grain boundary migration rates and temperatures) for specific materials and will give way at lower temperatures and higher strain rates to fracturing. When conditions do induce failure two possible modes are possible; a) propagation of a isolated single fractures which may link to produce a more extensive brecciated zone, or b) disaggregation of the aggregate by loss of cohesion between individual grains. Propagation to form single fractures appears to be responsible for the quartz veins which traverse the foliation in the Moine Thrust mylonites (Figure 3) while disaggregation along selected domains parallel to a mylonite foliation maybe more important to progressive localisation of slip in mylonites undergoing the transition to deformation by cataclastic flow (Figure 4).

### 3) Fault zone ageing and sealing : fluid pathways vs fluid flow barriers.

Changes in the porosity and permeability in a fault zone control the ability of a fault to focus fluid flow, allow cross fault flow and affect the mechanical response during later deformation events. This dual role of fault zones to act as both a high permeability pathways and as barriers to fluid movement is important to the understanding of the hydrodynamics and mechanics of fault zones as well as to the assessment of larger scale fluid migration. Processes of fault zone sealing or ageing are particularly important as these will determine the time period over which faults can act as high permeability windows. In addition to the juxtaposition of rock units with different sealing capacities, the mechanisms of fault zone sealing include:

1) Porosity Collapse -where reductions in permeability within fault zones are induced by changes in the pore volume, shape and size by grain refinement and rearrangement associated with frictional sliding and/or cataclasis, dissolution, alteration and plastic deformation of fragments and grains.
2) Cementation - where permeability reductions arise from the growth of new phases within and adjacent to the fault zone.

Microstructural analysis of the sealing processes of fault zones from ODP cores, sandstones recovered from hydrocarbon exploration wells and field based studies provide information on three important aspects on fault zone sealing (Knipe et al. 1991; Knipe 1992, 1993):

1) The combination of mechanisms responsible for sealing in faults developed in different lithologies at different depths.
2) Quantification of the changes in the porosity, permeability and pore geometry associated with different sealing mechanisms.
3) Evidence for the pattern of individual fault activity during the evolution of a fault array.

There is a range in the evolution of fault seals which depends upon the host rock lithology, the deformation mechanism path (i.e. the combination of deformation mechanisms which produce the final fault rock fabrics), and the potential of associated fluids to cause cementation. One of the interesting features to emerge from the microstructural analysis is the ability of fault zones developed in pure sandstone to generate low permeability, collapse seals seals, and the ability of small amounts (5-15%) of phyllosilicate in impure sandstone to generate smear seals (Knipe 1992). Figure 5 illustrates the sequence of microstructural evolution typical of sealing in cataclastic faults in a pure sandstones. The interlocking, low porosity aggregate which represents the end product of this self sealing process is generated by a combination of grain dissolution and precipitation.

Because of the juxtaposition of different rock types against an individual fault and the range of stress histories experienced at rupture nucleation areas, fault plane jogs and tip zones, both the fault rock evolution and the sealing properties will vary along individual faults. This, together with the complexities of the timing of displacement on individual faults in an array and fluid pressure evolution in different rock units, leads to fluid flow being controlled by the creation and destruction of complex 3D migration pathways (Knipe 1993). The influence of evolving fault plane geometries, displacement patterns, tip zone processes and fault rock evolution in controlling the juxtapositions and windows for fluid communication and diagenetic changes are discussed in more detail in Knipe (1993).

Microstructural analyses of the diagenetic histories within and outside fault zones recovered from hydrocarbon exploration well cores provides ideal material for assessing the links between deformation mechanisms, fault zone diagenesis and the influence of these processes on fluid flow on a larger scale. The combination of microstructural analyses of deformation and

Figure 5. Sequential evolution of a sealed cataclasite fault zone. The top SEM micrograph is of an undeformed sandstone (scale 3mm = 75μm); the central micrograph is of a partial sealed cataclasite (scale 3mm = 3μm), and the lower micrograph illustrates a low porosity, sealed gouge zone (scale 3mm = ~1μm).

diagenetic sequences together with the assessment of the propagation and geometrical evolution of faults provides an important route for the evaluation of; a) fault zone porosity/permeability histories, b) the evolution and distribution of fault seals; c) the timing of fault activity during burial and basin development; and d) identification of fluid migration pathways and the changing drainage patterns, which control the distribution and timing of open/closed compartments, within a large fault zone or a sedimentary basin. Integration of the microstructural analysis with the interpretation of fault plane geometry and displacement magnitude variations on fault planes (from the 3D seismic surveys) also allows an assessment of the distribution of seal types on fault planes. This integration aids the identification of "weak" and/or "leaky" seal areas on fault planes (see Knipe 1993).

**Conclusions.**

This short paper has highlighted the role that detailed microstructural analysis of natural fault zones can play in the assessment of the interaction of deformation and fluid processes.

Three aspects of the these interactions have been emphasised:

1) The interaction of fluid pressure cycles and damage accumulation in the rheology of cataclastic fault zones.
2) The role of 'creep void' development and linking in the initiation of instabilities near the 'brittle-ductile' transition.
3) The role of fault sealing processes in the creation of compartments in faulted areas.

**ACKNOWLEDGEMENTS.**

The author would like to express his thanks to the organisers of the meeting and to the USGS for support to attend the meeting.

**REFERENCES.**

Carter, N.L., Kronenburg, A.K., Ross,J.V. & Wiltschko, D.V. 1990 Control of fluids on deformation in rocks. In, Knipe, R.J. & Rutter, E.H. (Eds): Deformation Mechanisms, Rheology and Tectonics.Geol. Society Special Publication No 54 1-13.

Etheridge M.A., Wall, V.J., Cox, S.F. and Vernon, R.H. 1984. High fluid pressures during metamorphism and deformation: implications for mass transport and deformation mechanisms. J. Geophys. Res. 89, 4344-4358.

Knipe, R.J. 1990. Microstructural analysis and tectonic evolution in thrust systems: examples from the Assent region of the Moine Thrust Zone, Scotland. In: Deformation Processes in Minerals, Ceramics and Rocks. (): Barber, D.J. and Meredith, P.G. The Mineralogical Society Series. 1 228-258.

Knipe, R.J. 1992. Faulting processes and fault seal. In: Structural and Tectonic Modelling and it's application to petroleum Geology. Ed. R.M. Larsen et al. NPF, Special Publ. No 1. 325-343.

Knipe, R.J. 1993. The influence of fault zone processes and fault zone diagenesis on fluid flow. In diagenesis and Basin Development. Eds. A.D. Horbury and A.G. Robinson. AAPG Studies in Geology 36 135-154..

Knipe, R.J., Agar, S.M. and Prior, D.J. 1991. The microstructural evolution of fluid flow paths in semi-lithified sediments from subduction complexes. Phil. Trans. R. Soc. Lond. A, 335, 261-273.

Langseth, M. & Moore, J.C. Introduction to special section on the role of fluids in sediment accretion, deformation, diagenesis and metamorphism in subduction zones, J. Geophys Res. 95. 8737-8742.
Lloyd, G.E. and Knipe, R.J. 1992. Deformation mechanisms accommodating faulting of quartzite under upper crustal conditions. J. Struc. Geol. 14. 127-144.
Sibson, R.H. 1986. Earthquakes and rock deformation in crustal fault zones. Ann. Rev. Earth Planet. Sci., 149-175.
Sibson, R.H. 1989. Earthquakes as a structural process. J. Struc. Geol. 11. 1-14.
Sibson, R.H. 1992. Implications of fault -valve behaviour in rupture nucleation and recurrence. Tectonophysics. 211. 283-293.

# Remarks on the Mechanics and Kinetics of Permeability Changes in Rocks

Brian Evans, Matthias Imhoff, Gunter Siddiqi,
Massachusetts Institute of Technology
Department of Earth, Atmospheric, and Planetary Sciences
Cambridge, MA 02139-4307
e-mail: brievans@mit.edu

## Abstract

Dry thermal cracking at room pressure enhances permeability by an amount that depends primarily on the peak temperature, as observed earlier by others. There is wide scatter in the results, probably owing to inhomogeneities in the porosity of the starting samples. Thermal shock on nominally dry samples with superposed low or intermediate effective pressure (up to 100 MPa) enhanced the final permeability with the largest increases corresponding to the highest peak temperature. When an effective confining pressure of 200 MPa was superposed during thermal treatment on nominally dry samples, the permeability was not increased above the range found in intact samples. The effective pressure may suppress thermal cracking, or there may be some healing during the experiment, owing to traces of water trapped during pressurization. When pore fluids were present (200 MPa confining pressure, 100 MPa pore fluid pressure, 823 K), permeability remained low and healing apparently occured. Additional experiments at a broader range of temperatures and for longer times are needed to delineate the kinetics of permeability reductions and to understand the exact connection between permeability and crack-healing rates.

Earlier crack-healing experiments indicate that healing can occur either by a two-stage process involving formation of tube and bubbles, or by a single-stage process involving the healing of the crack as a single front. Those experiments indicated that the velocity of healing was a critical parameter in determing the mode of healing. New numerical treatments of healing driven by surface tension also became unstable when the velocity of the healing front became large, in agreement with the laboratory experiments.

## Introduction

### Mechanical Involvement of Fluids in Faulting

Widespread geologic evidence suggests that fluid pressures are sometimes elevated over a large fraction of the earth's crust, and syntectonic hydrofracturing events may be relatively common [*Fyfe et al.*, 1978; *Etheridge et al.*, 1984; *Walder and Nur*, 1984], perhaps at all metamorphic grades [*Ross and Lewis*, 1989]. High pore fluid pressure might also explain the apparent weakness of the San Andreas fault [*Byerlee*, 1990; *Rice*, 1992]. Recent friction experiments [*Blanpied et al.*, 1992] provide another example of the interplay between elevated pore fluid pressure, permeability, and rock strength.

In contrast, other observations require the presence of connected porosity and substantial flow of fluids at depth [e.g. *Sibson*, 1981; *Walder and Nur*, 1984; *Nur and Walder*, 1992]. But, if high pore pressures along faults are to be maintained over geologic time, either rock permeability must be rather low - on the order of 1 nanodarcy or $10^{-21}m^2$ [*Bredehoeft and Hanshaw*, 1968; *Nur and Walder*, 1992] or there must be an active source of overpressured fluids at depth [*Rice*, 1992].

The two sets of observations may be reconciled if permeability is regarded as a dynamic property. Thus, a quantitative knowledge of the rate of change of transport properties of rocks under various conditions of effective stress, temperature, and pore fluid chemistry is critical to understanding tectonic processes [*Gueguen et al.*, 1986; *Byerlee*, 1990; *Rice*, 1992; *Nur and Walder*, 1992].

### Mechanisms for Porosity Changes in Rocks

Mechanisms which both reduce and enhance porosity and pore networks have been recognized in field and experimental studies. Cracked rocks respond as a poroelastic material to small changes in stress, and the transport properties of the rocks are significantly affected [*Gangi*, 1979; *Walsh and Grossenbaugh*, 1979; *Tsang and Witherspoon*, 1981; *Carlson and Gangi*, 1985; *Brown*, 1987]. Many natural processes also involve inelastic deformation and modify pore geometry. Examples of network-producing mechanisms include fracturing owing to thermal or pressure cycling [*Heard and Page*, 1982; *Fredrich and Wong*, 1986; *Carlson et al.*, 1990], to volume changes during metamorphism [*Walther and Orville*, 1982], or to nonhydrostatic stresses [e.g. *Martin*, 1972; *Kranz et al.*, 1982; *Atkinson and Meredith*, 1987]; fluids may increase porosity by infiltrating rocks to establish textural equilibrium [*Watson and Brenan*, 1987; *Riley and Kohlstedt*, 1991]. Decreases in permeability might be caused by brittle pore collapse [*Zhang et al.*, 1990], pressure solution [*Elliot*, 1973; *Sprunt and Nur*, 1976, 1977a,b; *Engelder*, 1979, 1982, 1984; *Mosher*, 1980, 1987; *Mitra et al.*, 1984], crack healing, crack sealing [*Ramsay*, 1980; *Roedder*, 1984] and cementation [*Meyers*, 1974; *Houseknecht*, 1984; *Tada and Siever*, 1989]. All are common in sedimentary, metamorphic, and even igneous rocks.

Natural situations are often extremely complex, involving concurrent operation of several network reducing and producing mechanisms. For example, rocks might undergo metamorphism while minerals were exported or imported through the pore fluids. The influx of new minerals into a rock might fill grain interstices, form an interlocking texture, and indurate sediments [*Meyers*, 1974]. Simultaneously, dilatant cracking under load and pressure solution might occur, possibly in heterogeneous fashion.

Each mechanism has particular driving forces: for example, local non-hydrostatic stresses drive deformation at asperities and pores; reduction of interface tension drives crack healing and textural equilibration; chemical potentials drive metamorphism, dissolution into undersaturated liquids, or precipitation from oversaturated fluids. To identify the dominant mechanism, kinetics laws need to be developed for each process, and interactions between the mechanisms delineated. Rates could then be compared, given knowledge of the stresses, temperatures, fluid chemistry, pore pressures, pore microstructure, and mineralogy.

In this progress report, we describe experiments and modeling designed to explore the mechanics and kinetics of two processes which alter permeability: thermal cracking and crack healing. In the experiments we measured permeability before and after heat-treatments at effective pressures up to 200 MPa. Most experiments were done under dry, vented conditions, but some were done with an aqueous pore fluid present. The modeling effort involved finite difference computations of the change of shape of pores under the influence of surface tension forces as occurs during crack healing. Work on both parts of the project is continuing, but some trends are already apparent.

## Evolution of Permeability Owing to Thermal Cracking

Previous workers have investigated thermal cracking in granite [*Bauer and Johnson*, 1979; *Fredrich and Wong*, 1986; *Wang et al.*, 1989; *Carlson et al.*, 1990], and the effect of thermal cracking at low pressure on permeability [*Darot et al.*, 1992]. To add data from a new material under different conditions of pressure and temperature, we investigated the effect of thermal cracking on the evolution of permeability in Sioux quartzite. The study also forms a baseline for new experiments on the kinetics of permeability changes in quartzites.

We cracked the specimens by heating them to peak temperatures ranging up to 830K, at effective pressures up to 200 MPa, with constant heating/cooling rates, and with and without pore fluids. Subsequently, permeability was measured using the transient flow method at several effective pressures. Some samples were subjected to a four-stage process which included 1.) heat treatment at room pressure, 2.) permeability measurements at room temperature (actually about 45°C), but elevated pressures, 3.) hydrothermal treatment with both confining pressure and pore fluids present, and, finally, 4.) a second set of room temperature permeability measurements. Future experiments will include measurements of permeability at high temperature and pressure.

The starting material, Sioux quartzite, is a homogenous, well-indurated, sedimentary quartzite, dominantly composed of quartz (99%) with some oxide accessory minerals and about 0.6% porosity. The samples were ground to be cylinders approximately 19 mm in diameter and 25 mm long. The permeameter we used [*Bernabe*, 1987a] was capable of measuring permeabilities of $1\times10^{-23}$ $m^{-2}$, utilizing the pulse decay or transient method [*Brace et al.*, 1968]. The confining medium was kerosene and distilled water was used as the pore fluid.

*Bernabe* [1987b] studied the effect of loading path on permeability and concluded that pressure history does influence permeability in freshly cored samples. His observations led us to adopt the following test protocol: after thermal equilibrium is attained, confining pressure is raised to 1 MPa. The system is flushed of any remaining visible air bubbles and (rarely) contaminants. Confining pressure and pore fluid pressure were then gradually raised in small steps to approximately 170 MPa and 16 MPa, respectively. After the pressures equilibrated, we determined the permeability at a unique combination of confining pressure and pore fluid pressure. Keeping pore fluid pressure constant, we lowered confining pressure in steps to 60 MPa and then raised confining pressure back to the initial value of 170 MPa in a series of steps. During each step we performed a permeability measurement. Up to three cycles of pressurization and depressurization were performed on some samples.

Vented thermal-cracking experiments were performed on cylindrical samples of the same dimensions as for the permeability tests. For most experiments, a confining pressure of either 0.1, 100 or 200 MPa was applied first, then temperature was raised to 298, 683, 783, 830K at a constant heating rate of 5 K/min. Once at peak, temperature was held constant for 30 min. and then decreased at the same rate. The sample was retrieved and rejacketed, and a series of permeability tests performed. For some experiments, a second thermal-cracking treatment was done, but now with added pore fluid (unvented). Finally, in order to test the effect of heating rate, we varied heating rate in a few tests at 683 K and confining pressure of 100 MPa.

For microstructural observations crack sections were cut perpendicular to the axis of the cylinders, polished and ion-milled. We counted the crack intersections with a parallel test grid following *Fredrich and Wong* [1986] under an optical microscope. The grid consisted of 20 lines, equally spaced 50.1 μm apart, of 1.02 mm length each. The total area covered in each sample was 5 $mm^2$. The specific crack surface area ($mm^2/mm^3$) is then simply twice the value of the crack density [*Underwood*, 1970].

## Results of the Experiments

To establish a baseline, we measured permeabilities of a number of untreated samples over effective pressures from 36 MPa to 162 MPa. Permeabilities range from $58\times10^{-21}$ $m^2$ (58 nanodarcy) at effective pressures of app. 40 MPa, to $2\times10^{-21}$ $m^2$ at 162 MPa (Figure 1). Notice that five of the eight samples have a permeability ranging from 20 to $55\times10^{-21}$ $m^2$ at high and low effective pressures respectively, while three other samples have values of 2

to $9\times10^{-21}$ $m^2$. The error bars reflect uncertainties in the individual permeability tests. The mean number of crack intersections per unit length, $P_L$, in the untreated rocks, is $6.08\pm2.12$ $mm^{-1}$, corresponding to a specific crack surface area of 12.16 $mm^2/mm^3$. Most cracks run along grain boundaries; pore porosity is minor and concentrated at grain junctions.

The crack experiments have three adjustable parameters: heating rate, peak temperature and confining pressure during heating and cooling; each of these were varied within limits. Thermal cracking at room pressure increases permeability by up to a factor of three, the largest increases corresponding to highest peak temperatures. The most permeable untreated samples have permeabilities ranging from 50 to $80\times10^{-21}$ $m^2$, but the samples cracked at room temperature showed permeabilities as large as $190\times10^{-21}$ $m^2$. In the heat-treated samples, grain boundaries are extensively cracked and crack widths may have increased. The specific surface area of one 830 K specimen was 16.70 $mm^2/mm^3$.

To observe effects due to heating rate, we cracked four samples at a confining pressure of 100 MPa and peak temperature of 683 K, but at heating/cooling rates of 1, 5, 10 K/min respectively. There was no significant correlation of permeability with heating rate. The first 5 K/min sample was more permeable than all the others, but was not as permeable on subsequent pressure cycles. We performed a duplicate 5 K/min experiment which did not have elevated permeability. We conclude that heating rates of less than 10 K/min have no discernible effect on permeability. In all subsequent experiments we heated and cooled at 5 K/min.

Superposing intermediate confining pressures of 100 MPa during heat treatment results in significant permeability increases only for those specimens heated to 830 K. Measured values ranged from 40 - $85\times10^{-21}$ $m^2$ at high effective pressures to 90 - $140\times10^{-21}$ $m^2$ at low effective pressures. One sample was very permeable, compared to the others which underwent the same treatment: 105 and $480\times10^{-21}$ $m^2$ at high and low effective pressures.

In general, permeabilities for samples heated to 683 and 783 K tend to have lower permeabilities than those heated to 830 K (Figure 2). The first cycle of one specimen taken to 683 K and 100 MPa confining pressure did show a significant increase above background, but this increase did not persist in a second and third pressure cycle. A repeat experiment did not duplicate the increased permeability of the first. The average number of crack intersections per unit length was $7.62\pm1.98$ $mm^{-1}$ for the 783 K experiments.

Heat-treatments of nominally dry samples at the highest confining pressures (200 MPa) did not result in increased permeability (Figure 3). In fact, the permeability of those samples is closer to the lower bound of the untreated samples: values range from $1\times10^{-21}$ $m^2$ at high effective pressures to $20\times10^{-21}$ $m^2$ at low effective pressure. It appears that a combination of elevated temperature and high confining pressure can keep permeabilities low (or possibly even decrease permeability). The cracks in these samples do not often run along entire grain boundaries, and there are more bubble trains than in the untreated sample. The specific crack surface area of one sample heated to 830K is 10.86 $mm^2/mm^3$.

Some preliminary tests under hydrothermal conditions have also been done. The samples that were cracked at 830 K and 100 MPa confining pressure were the starting material for a second hydrothermal crack treatment at 200 MPa confining pressure and pore fluid pressure of 100 MPa. Figure 4 shows the permeability after the first, nominally dry thermal treatment of 830 K, 100 MPa confining pressure and after the second hydrothermal treatment. Two samples (P1003 and P1017) exhibit a small drop in permeability, while the third sample (P1011) drops by a factor of 3. In the former two experiments, pore pressure leaks developed, possibly influencing the results. The thin sections show extensive healing structures including pinched off cracks, fluid inclusions, and bubble trains. The specific crack surface area of P1017 is 8.84 $mm^2/mm^3$.

## Discussion

Acoustic emission rate increases with heating rate [*Yong and Wang,* 1980], but there is conflicting evidence of the effect of heating rate on the total amount of crack porosity [*Thirumalai and Demou,* 1973; *Yong and Wang,* 1980]. Although there is some scatter to the data, these Sioux quartzite experiments agree with experiments of *Yong and Wang* [1980] and *Fredrich and Wong* [1986], and suggest that peak temperature is more important in inducing cracks than are changes in heating rate. A heat flow calculation using the solution for a cylinder heated from outside [*Williamson and Adams,* 1919] suggests that the temperature gradients during heating are at most a few degrees per millimeter. Apparently the stresses generated by this gradient alone are much too low to generate a noticeable increase in the crack densities.

Increasing pressure suppresses the formation of thermal cracks. Intuitively, one expects that increasing pressure will increase the critical stress intensity factor necessary to propagate flaws. We quantitatively estimated the stress intensity factors for a square quartz inclusion in a homogeneous matrix using *Fredrich and Wong's* [1986] fracture mechanics model modified to include the effect of a superposed confining pressure [*Cartwright and Rooke,* 1979]. For a flaw oriented along the grain boundary of the square inclusion, the calculation predicts that a confining pressure of 200 MPa will lower the stress concentration factor below the critical stress intensity factor for all flaw sizes, but that some short flaws may propagate if the pressure is lowered to 100 MPa.

## Summary

Dry thermal cracking at room pressure enhances permeability by an amount that depends primarily on the peak temperature, as observed earlier by *Fredrich and Wong* [1986], but there is wide scatter in the results, probably owing, in part, to inhomogeneities in the porosity of the starting samples. Thermal shock on nominally dry samples with superposed low or intermediate effective pressure (up to 100 MPa effective pressure) enhances the final permeability with the largest increases corresponding to the highest peak temperature. When pore fluids are present, healing occurs, and in one experiment, permeability was reduced by a factor of about 3. The hydrothermal conditions were 200

MPa confining pressure, 100 MPa pore fluid pressure, 823K up to two hours at peak temperature. Crack-healing processes apparently modify pore geometry, and can be observed by optical and scanning electron microscopy. The crack surface area per unit volume drops from 14 to 9 $mm^{-1}$.

When an effective confining pressure of 200 MPa was superposed during thermal treatment on nominally dry samples, the permeability was not increased above the range found in intact samples. The effective pressure may suppress thermal cracking, as is suggested by calculations of the magnitude of thermal stresses arising at the boundaries. But, there is also microstructural evidence of healing during the experiment, suggesting that trace amounts of water are trapped in the sample upon application of the pressure.

## Evolution of Porosity Owing to Capillary Forces

Most rocks contain healed fractures, often marked by fluid inclusion planes [*Simmons and Richter*, 1976; *Sprunt and Nur*, 1979; *Hollister and Crawford* , 1981; *Roedder*, 1984]. If the crack-filling mineral is the same as the host, the cracks are said to be healed. Cement may come from local sources, or outside the formation, perhaps entering by advection in a supersaturated fluid [e.g. *Simmons and Richter*, 1976; *Fisher and Brantley*, 1992]. Even with stagnant fluids, crack healing is extensive in quartz at low temperature [*Smith and Evans*, 1984; *Brantley et al.*, 1989].

When gradients in boundary curvature exist, interface tension drives material redistribution via surface diffusion, lattice diffusion, or transport through pores [*Blakely*, 1973]. In rocks, fluids often provide a fast transport path. Then, interface tension between grain boundaries and fluid/mineral interfaces will cause intergranular porosity to evolve to tubules along grain edges, isolated fluid pockets, or sheet-like pores along grain faces, depending on fluid chemistry and interface energy [*Beere*, 1975; *Bulau et al.*, 1979; *Watson and Brenan*, 1987; *Watson et al.*, 1991; *Lee et al.*, 1991; *Laporte and Watson*, 1991; *Holness*, 1991]. Examples of studies of grain boundary healing are *Hickman and Evans* [1991] in halite and *Olgaard and Fitzgerald*, [unpublished work, 1994] in calcite. Intragranular porosity will evolve from cracks to inclusion planes or to a single pore depending on healing velocity [*Roedder*, 1984; *Hickman and Evans*, unpublished work, 1994]. Healing rates are related to temperature [*Shelton and Orville*, 1980; *Smith and Evans*, 1984; *Bodnar and Sterner*, 1985, 1987; *Wanamaker and Evans*, 1985; *Brantley et al.*, 1989], fluid chemistry [*Brantley and Voigt*, 1989], and crack size [*Hickman and Evans*, 1987].

Laboratory experiments and many, but not all, observations of naturally healed rocks indicate that healing often proceeds via a two-stage process involving the production of cylindrical voids near the edge of the healing crack tip, and the subsequent evolution of the cylindrical voids into spherical fluid inclusions. Based on treatments of the instability of a cylindrical pore [*Nichols and Mullins* 1965a,b; *Nichols*, 1976], *A. G. Evans and R. A. Charles* [1977] suggested that the healing distance was given by

$$\frac{a-r}{a} = E_m \left( \frac{t \cdot exp\left(\frac{-Q}{RT}\right)}{T \cdot ab_r^w} \right)^n \tag{1}$$

where $r$ is the distance from the crack center to the healing tip, $a$ is the original crack length, $t$ is the elapsed time, $E_m$ is a constant, $Q$ is the activation energy for transport, $R$ is the gas constant, $a$ and $b_r$ are the semi-major and semi-minor axes of the original crack, and $w$ and $n$ are constants which depend on geometry of cylinder formation. Eq. 1, as given, is a slight modification of *A. G. Evans and Charles* [1977] to apply to cracks with elliptical cross-section [*Hickman and Evans*, 1987].

Similar systematics might apply to crack healing in fluid-saturated rock masses with the important difference that dissolution, electrolyte diffusion, and precipitation are transport steps. The treatment assumes that the cracks are elliptical, that crack surfaces are not forced together except at the tips, that fluids are not transported into or out of the solid, and, of course that crack healing proceeds by tube and bubble formation. Pore fluid pressure remains constant.

Observations of naturally healed rocks using optical, scanning electron, and cathodoluminescence microscopy sometimes indicate that healing has occurred, but fluid inclusions are not apparent along the trace of the healed fracture [*Simmons and Richter*, 1976; *Sprunt and Nur*, 1979; *Hollister and Crawford*, 1981; *Roedder*, 1984]. This second mode of healing, which we call continuous healing, was observed recently in a set of crack-healing experiments in the halite/water system [*Hickman and Evans,* 1994]. The mechanics and kinetics of this mode of healing are important because Eq. 1 may no longer be valid, and because the transition in mode might provide a way to constrain healing rates in natural rocks.

## Stability Analyses

To investigate the transition from continuous-front healing to tube-and-bubble behavior we performed a set of finite difference calculations on tubes and cracks of different geometry, assuming the following. The system was closed to the addition of new solid. Temperature was constant. All material and interface properties are isotropic; interface energy depends only on the radius of curvature of the surface. The crack or tube is a cylindrical, parabolic surface, the thickness of which can be expressed as a gently varying function of distance from the tip. And, importantly, material transport takes place by surface diffusion alone. Some results are shown in Figures 5 and 6.

In one set of simulations, we investigated the stability of a tube with sinusoidal perturbations along the axis of the cylinder. The results agree with the analytical solution of *Nichols and Mullins* [1976a,b] for a cylinder with very small perturbations. The tube is

more likely to be unstable and pinch into two cylinders with closed ends when the perturbations are large, but surprisingly large constrictions in the radius of the tube resist pinching off. Simulating the healing of an ellipsoidal crack is a taxing calculation. An ellipsoidal prism with the ratio of major to minor semi-axis of 0.94 was unstable, while those prisms which were less eccentric were stable, mimicking at least one aspect of the natural and experimental observations. Simulations for more complex and realistic shapes are very time-consuming, so that it is important to make at least some simplifying assumptions. One important extension on which we are working is to allow pore fluid transport under realistic conditions.

If under- or over-saturated fluids were to circulate through the formation, rates might be quite different. Crack-healing process with local sources will dominate when length scales associated with flowing fluids are smaller than those associated with diffusion [e.g. *Fisher and Brantley*, 1992]. Heterogeneity in stress, plastic strain, surface energy, or grain size might also affect healing rates. Since the gradients in chemical potential arising from capillary forces are always present, the healing rates calculated from Eq. 1 may be considered lower bounds. Crack healing may also be important in establishing a threshold rate for stable crack growth [*Atkinson*, 1984].

**Discussion : Solution Transport Processes and Permeability**

Several studies show that permeability can be strongly affected by solution-transport on laboratory time scales: In triaxial creep experiments on granite and sandstone, *Kranz and Blacic* [1984] observed permeability which increased, but then decreased. Strain and permeability data suggested that dilatancy owing to microcracking competed with permeability reductions by dissolution-precipitation. Dramatic permeability changes also occurred in tuff and granite over periods of weeks, as pore fluids flowed down a temperature gradient [*Morrow et al.*, 1981; *Moore et al.*, 1983; *Vaughan et al.*, 1986]. There was significant silica mobility and growth of calcium-rich minerals. *Blanpied et al.* [1992] demonstrated healing which hydraulically isolated a simulated fault from the external pore pressure reservoir. Once the fault sealed, gouge compaction elevated the pore-pressure, allowing slip to occur at stresses well below those for drained conditions.

Relating pore structure to transport properties is a durable problem [*Dullien*, 1979; *Katz and Thompson*, 1986; *Berryman and Blair*, 1987; *Brown and Scholz*, 1986; *Brown*, 1987; *Doyen*, 1988; *Pyrak-Nolte and Cook*, 1988; *Gueguen and Deines*, 1989; *David et al.*, 1990]. One key issue has been identifying the critical characteristics of pore shape, pore dimensions, or connectivity. For example, permeability in cracked solids is often modeled using the equivalent channel model:

$$k = \left(\frac{\Phi}{S_V}\right)^2 \frac{1}{3F} \qquad (2)$$

where $\Phi$ is the total porosity, $S_V$ is the surface area of pores per unit volume of solid, and $F$ is a formation factor, such that $F=t^2/\Phi$, and $t^2$ is the tortuosity [*Walsh and Brace*, 1984; *Paterson*, 1984]. Thus, pore structure is characterized by three parameters, $F$, $S_V$, *and* $t^2$. Although successful in modeling some aspects of permeability, the equivalent channel model contains no information about number distribution of size and shape or about connectivity, except in the single constant, $t^2$.

*Brantley et al.* [1989] calculated the effect of crack healing on permeability using the equivalent channel model for three hypothetical microcrack populations assuming that all cracks are identical elliptical tubes, parallel to the fluid pressure gradient; and that the regression distance is given by Eq.1 with n = 0.4. When the crack apertures are small, healing is very rapid and the permeability decreases precipitously, but when crack apertures are large, i.e. fewer, but fatter cracks, permeability is affected only slowly. Clearly the model is only a cartoon of the real process; it completely neglects an extremely important issue: how is the connectivity of the pore space modified? This question is, of course, central to the determination of permeability [*Madden*, 1983; *Long and Witherspoon*, 1985; *Gueguen and Dienes*, 1989; *David et al.*, 1990].

The actual geometry of pores in granular sedimentary rocks [*Thompson*, 1991], of the topography of rock fractures [*Brown and Scholz*, 1986; *Brown*, 1987], and of cracks in crystalline rocks [*Wong et al.*, 1989] is that of a highly connected network with broad distribution of amplitudes, sizes, and aspect ratios. *Wong et al.* [1989] show a frequency distribution of crack aperture and crack surface density for Rutland quartzite and for Westerly granite . The exact statistics and average values to be found in a particular rock depend on its strain, temperature, pressure, and metamorphic history [*Nur and Simmons*, 1970; *Padovani et al.*, 1982; *Doyen*, 1987; *Wong et al.*, 1989; *Thompson*, 1991]. When textural forces dominate, as might occur at depth in a tectonically inactive area, one expects the frequency distribution to be altered from a broad distribution with many small cracks to a more narrow distribution of larger cracks. Eventually the network should evolve either to a connected network of tubes along grain boundaries or isolated pores depending on mineralogy and pore fluid chemistry [*Watson and Brenan*, 1987; *Lee et al.*, 1991; *Holness*, 1992].

When broad distributions of pore size and aperture are present, percolation and network models may be more useful than the equivalent channel model in relating permeability to pore structure [*David*, 1993]. Using a percolation model assuming a uniform number distribution of crack lengths and apertures, *Gavrilenko and Gueguen* [1993] suggest that permeability may written as:

$$k = \frac{4\pi}{15} N_f w_o^3 \bar{c}^2 F(X)(1-S) \tag{3}$$

where $N_f$ is the crack number density, $\bar{c}$ is the mean crack length, $F(X)$ is the fraction of cracks which contribute to flow [given by *Gueguen and Dienes*, 1989), $X$ is the probability that two cracks intersect, an exponential function of $N_f \bar{c}^3$, $S$ is a tortuosity

factor. *David et al.* [1990] suggest that the nature of the distribution function is an important factor in determining permeability. Using renormalization group theory, *Madden* [1983] predicted that crack connectivity is most critically determined by the number density of cracks times the square of the average crack length. If the porosity is composed of a connected set of grain-edge tubes, then permeability should be roughly proportional to the mean radius of the tube to the fourth power [see review by *David et al.*, 1990].

**Summary**

Additional laboratory and numerical experiments at a broader range of conditions are needed to delineate the kinetics of permeability reductions and to understand the exact connection between permeability and crack-healing rates, but it seems clear that substantial reductions in permeability can be achieved in relatively short periods of time, even when no new minerals are added to the rock as cement. Because the healing models discussed here incorporate the driving force owing to interface tension only, the models will predict lower bounds of the healing rates. Comparison of these models with others that predict crack extension rates and crack healing rates for other driving forces could be developed to yield multi-dimensional maps of crack extension and healing in regions in $P_c$, $P_f$, T, and $\sigma$ space.

## Acknowledgments

Gunter Siddiqi was supported by a fellowship generously provided by the Royal Dutch Shell Company. The Earth Resources Laboratory at M.I.T. allowed us to use their massively parallel computer. Thanks also go to the organizers of the USGS conference.

## References

Atkinson, B.K., Sub-critical crack growth in geological materials, *J. Geophys. Res., 89,* 4077-4113, 1984.

Atkinson, B.K. and P.G. Meredith, Experimental fracture mechanics data for rocks and minerals, In: *Fracture Mechanics of Rock*, B.K. Atkinson (Ed.), 477-527, 1987.

Bauer, S. J., and B. Johnson, Effects of slow uniform heating on the physical properties of the Westerly and Charcoal granites, *Proc. U.S. Symposium Rock Mech., 20,* 7-18, 1979.

Beere, W., A unifying theory of the stability of penetration liquid phases and sintering pores, *Acta Metall., 23,* 131-138, 1975.

Bernabe, Y., A wide range permeameter for use in rock physics, *Int. J. Rock Mech. Min. Sci. Geomech. Abstr., 24,* 309-15, 1987a.

Bernabe, Y., The effective pressure law for permeability during pore pressure and confining pressure cycling of several crystalline rocks. *J. Geophys. Res., 92 (1)*, 649-657, 1987b.

Bernabe, Y., W.F. Brace, and B. Evans, Permeability, porosity, and pore geometry of hot-pressed calcite. *Mech. Mater. 1*, 173-183, 1982.

Berryman, J. G., and S. C. Blair, Kozeny-Carman relations and image processing methods for estimating Darcy's constant, *J. Appl. Phys. 62*, 2221-2228, 1987.

Blakely, J.M., *Introduction to the properties of crystal surfaces*, Pergamon Press, Oxford, 1973.

Blanpied, M.L., D.A. Lockner, and J.D. Byerlee, An earthquake mechanism based on rapid sealing of faults, *Nature, 358*, 574-576, 1992.

Bodnar, R. J., and S. M. Sterner, Synthetic fluid inclusions in natural quartz, II. Application to PVT studies, *Geochimica et Cosmochimica Acta, 49*, 1855-1859, 1985.

Brace, W.F., J.B. Walsh and W.T. Frangos, Permeability of granite under high pressure, *J. Geophys. Res., 73*, 2225-2236, 1968.

Brantley, S.L., and D. Voight, Fluids in metamorphic rocks: Effects of fluid chemistry on quartz microcrack healing, in *Water-Rock Interaction WRI-6: Proc. of 6th Inter. Sym. on Water Rock Interaction*, Malvern, edited by D.L. Miles, pp. 113-116, A.A. Balkema, Rotterdam, 1989.

Brantley, S. L., B. Evans, S. H. Hickman, and D. A. Crerar, Healing of microcracks in quartz: Implication for fluid flow, *Geology, 18*, 136-139, 1989.

Bredehoeft, J.D. and B.B. Hanshaw, On the maintenance of anomalous fluid pressures, I. Thick sedimentary sequences, *Geol. Soc. Am. Bull., 79*, 1097-1106, 1968.

Brown, S. R., Fluid flow through rock joints: The effect of surface roughness, *J. Geophys. Res., 92*, 1337-1347, 1987.

Brown, S. R. and C. H. Scholz, Closure of rock joints, *J. Geophys. Res., 91*, 4939-4948, 1986.

Bulau, J. R., H. S. Waff, and J. A. Tyburczy, Mechanical and thermodynamic constraints on fluid distribution in partial melts, *J. Geophys. Res., 84*, 6102-6108, 1979.

Byerlee, J.D. Friction, overpressure and fault normal compression, *Geophys. Res. Lett. 17*, 2109-2112, 1990.

Carlson, R. L., and A. F. Gangi, Effect of cracks on the pressure dependence of P wave velocities in crystalline rocks, *J. Geophys. Res., 90,* 8675-8684, 1985.

Carlson, S.R., M. Wu, and H.F. Wang, Micromechanical modeling of thermal cracking in granite, in *The Brittle-Ductile Transition in Rocks, The Heard Volume, Geophys. Monogr. Ser., Vol. 56,* edited by A.G. Duba et al., pp. 37-48, AGU, Washington, D. C., 1990.

Cartwright D.J. and D.P. Rooke, Green's functions in fracture mechanics, In: *Fracture Mechanics, Current Status, Future Prospects,* R.A. Smith (Ed.), Pergamon, New York, 91-121, 1979.

Darot, M., Y. Gueguen and M.-L. Baratin, Permeability of thermally cracked granite, *Geophys. Res. Lett., 19,* 869-872, 1992.

David, C., Geometry of flow paths for fluid transport in rocks, *J. Geophys. Res., 98,* 12,267-12,278, 1993.

David, C., Y. Gueguen, and G. Pampoukis, Effective medium theory and network theory applied to the transport properties of rock, *J. Geophys. Res., 95,* 6993-7006, 1990.

Doyen, P. M., Permeability, conductivity and pore geometry of sandstone, *J. Geophys. Res., 93,* 7729-7740, 1988.

Dullien, F. A. L., *Porous Media, Fluid Transport and Pore Structure,* Academic Press, New York, 1979.

Elliot, D., Diffusion flow law in metamorphic rocks, *Geol. Soc. Amer. Bull., 84,* 2645-2664, 1973.

Engelder, T., Mechanisms for strain within the Upper Devonian clastic sequence of the Appalachian plateau, Western New York, *Amer. J. Sci, 279,* 527-542, 1979.

Engelder, T., A natural example of the simultaneous operation of free-face dissolution and pressure solution, *Geochim. Cosmochim. Acta, 46,* 69-74, 1982.

Engelder, T., The role of pore water circulation during the deformation of foreland fold and thrust belts, *J. Geophys. Res. 89,* 4319-4326, 1984.

Etheridge, M.A., V.J. Wall, S.F. Cox and R.H. Vernon, High fluid pressures during regional metamorphism and deformation: implications for mass transport and deformation mechanisms, *J. Geophys. Res., 89,* 4344-4358, 1984.

Evans, A.G., and E.A. Charles, Strength recovery by diffusive crack healing, *Acta Metall., 25,* 919-927, 1977.

Fertl, W.H., Abnormal Formation Pressures: Implications to Exploration, Drilling, and Production of Oil and Gas Resources, *Developments in Petroleum Science, 2,* Elsevier, 1976.

Fisher, D. M. and S. L. Brantley, Models of quartz overgrowth and vein formation: Deformation and episodic fluid flow in an ancient subduction zone, *J. Geophys. Res.,* 97, 20,043-20,061, 1992.

Fredrich, J. T., and T.-f. Wong, Micromechanics of thermally induced cracking in three crustal rocks, *J. Geophys. Res., 91,* 12743-12764, 1986.

Fyfe, W.S., N. J. Price, and A. B. Thompson, *Fluids in the Earth's Crust, Developments in Geochemistry 1,* Elsevier Scientific, New York, 1978.

Gangi, A. F., Variation of whole and fractured porous rock permeability with confining pressure, *Int. J. Rock Mech. Min. Sci., 15,* 249-257, 1979.

Gavrilenko, P. and Gueguen, Y. Fluid overpressures and pressure solution in the crust, *Tectonophysics, 217,* 91-110, 1993.

Gueguen, Y., and J. Dienes, Transport Properties of Rocks from Statistics and Percolation, *Math. Geol., 21,*1-13, 1989.

Gueguen, Y., C. David and M. Darot, Models and Time Constants for Permeability Evolution, *Geophys. Res. Lett., 13,* 460-463, 1986.

Heard, H.C., and L. Page, Elastic moduli, thermal expansion and inferred permeability of two granites to 350°C and 55 MPa, *J. Geophys. Res., 87,* 9340-9348, 1982.

Hickman, S. H., and B. Evans, Influence of geometry on crack healing rate in calcite, *Phys. Chem. Minerals, 15,* 91-102, 1987.

Hickman, S. H. and Evans, B., Experimental Pressure Solution in Halite, 1: The effect of grain/interphase boundary structure, *J. Geological Soc. London, 148,* 549-560, 1991.

Hollister, L. S. and Crawford, M. L. (ed.), *Short Course in Fluid Inclusions: Applications to Petrology,* Mineralogical Assoc. Canada, Calgary, Canada, 1981.

Holness, M. B., Equilibrium dihedral angles in the system quartz-$CO_2$-$H_2O$-NaCl at 800°C and 1-15 kbar: the effects of pressure and fluid composition on the permeability of quartzites, *Earth Planet. Sci. Lett. 114,* 171-184, 1992.

Houseknecht, D. W., Influence of grain size and temperature on intergranular pressure solution, quartz cementation, and porosity in a quartzose sandstone, *J. Sediment. Petrol. 54,* 348-361, 1984.

Katz, A. J. and A. H., Thompson, Quantitative prediction of permeability in porous rock, *Phys. Rev., B 34,* 8179-8181, 1986.

Kranz, R. L., and J. D. Blacic, Permeability changes during time-dependent deformation of silicate rock, *Geophys. Res. Lett., 11,* 975-978, 1984.

Kranz, R. L., W. J. Harris, and N. L. Carter, Static fatigue of granite at 200°C, *Geophys. Res. Lett., 9,* 1-4, 1982.

Laporte, D. and E. B. Watson, Direct observation of near-equilibrium pore geometry in synthetic quartzites at 600-800°C and 2-10.5 kbar, *J. Geophys. Res. 99,* 873-878, 1991.

Lee, V. W., S. J. Mackwell and S. L. Brantley, The effect of fluid chemistry on wetting textures in novaculite, *J. Geophys. Res., 96,* 10,023-10,037, 1991.

Long, J.C., and P.A. Witherspoon, The relationship of the degree of interconnection to permeability in fracture networks, *J. Geophys. Res., 90,* 3087-3098, 1985.

Madden, T.R., Microcrack connectivity in rocks: a renormalization group approach to the critical phenomena of conduction and and failure in crystalline rocks, *J. Geophys. Res., 88,* 585-592, 1983.

Martin, R.J., Time-dependent crack growth in quartz and its application to the creep of rocks, *J. Geophys. Res., 77,* 1406-1419, 1972.

Meyers, W. J., Carbonate cement stratigraphy of the Lake Valley Formation (Mississippian) Sacramento Mountains, New Mexico, *Sed. Pet. 44,* 837-861, 1974.

Mitra, G., W.A. Yonkee, and D.J. Gentry, Solution cleavage and its relationship to major structures in the Idaho-Utah-Wyoming thrust belt, *Geology, 12,* 354-358, 1984.

Moore, D.E., C.A. Morrow, and J.D. Byerlee, Chemical reactions accompanying fluid flow through granite held in a temperature gradient, *Geochim. Cosmochim. Acta., 47,* 445-453,1983.

Morrow, C., D. Lockner, D. Moore and J. Byerlee, Permeability of granite in a temperature gradient, *J. Geophys. Res., 86,* 3002-3008, 1981.

Mosher, S., Pressure solution deformation of conglomerates in shear zones, Narraganset Basin, Rhode Island, *J. Struct. Geol. 2,* 219-226, 1980.

Mosher, S., Pressure-solution deformation of Purgatory Conglomerate, Rhode Island (U.S.A.): quantification of volume change, real strains and sedimentary shape factor, *J. Struct. Geol. 9,* 221-232, 1987.

Nekut, A., J.E.P. Connerney, and A.F. Kuckes, Deep crustal electrical conductivity: Evidence for water in the lower crust. *Geophys. Res. Lett., 4*, 239-242, 1977.

Neuzil, C.E., Groundwater flow in low-permeability environments, *Water Resources Research, 22,* 1163-1195, 1986.

Nichols, F.A., Spheroidization of rod-shaped particles of finite length, *J. Mater. Sci, 11,* 1077-1082, 1976.

Nichols, F.A., and W.W. Mullins, Morphological changes of a surface of revolution due to a capillarity-induced surface diffusion, *J.Appl. Phys., 36,* 1826-1835, 1965a.

Nichols, F.A., and W.W. Mullins, Surface (interface) and volume diffusion contributions to morphological changes driven by capillarity, *Trans. AIME, 233*, 194-1948, 1965b.

Nur, A., and G. Simmons, The origin of small cracks in igneous rocks, *Int. J. Rock Mech. Min. Sci., 7,* 307-314, 1970.

Nur, A., and J. Walder, Hydraulic pulses in the earth's crust, in *Fault Mechanics and Transport Properties in Rocks: A Festschrift in Honor of W.F. Brace*, edited by B. Evans and T.-f. Wong, pp. 461-473, 1992.

Padovani, E., S. B. Shirey, and G. Simmons, Characteristics of microcracks in amphibolite and granulite facies grade rocks from southeastern Pennsylvania, *J. Geophys. Res., 87,* 8605-8630, 1982.

Paterson, M. S., The equivalent channel model for permeability and resistivity in fluid saturated rock - A re-appraisal, *Mech. Mat., 2,* 345-352, 1983.

Pyrak-Nolte, L. J. and N. G. W. Cook, Fluid percolation through single fractures, *Geophys. Res. Lett. 15,* 1247-1250, 1988.

Ramsay, J.G.,The crack-seal mechanism of rock deformation, *Nature 284,* 135-139, 1980.

Rice, J.R., Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault, in *Fault Mechanics and Transport Properties in Rocks: A Festschrift in Honor of W.F. Brace,* edited by B. Evans and T.-f. Wong, pp. 475-503, 1992.

Riley, G.N., Jr. and D.L. Kohlstedt, Kinetics of melt migration in upper mantle-type rocks, *Earth Plan. Sci. Lett., 105*, 500-521, 1991.

Roedder, E., Fluid Inclusions, *Reviews in Mineralogy, 12*, Mineralogical Society of America, Washington, DC, pp. 70-77, 1984.

Ross, J.V., and P.D. Lewis, Brittle-ductile transition: Semi-brittle behavior, *Tectonophysics, 167,* 75-79, 1989.

Shankland, T.J., R.J. O'Connell, and H.S. Waff, Geophysical constraints on partial melt in the upper mantle, *Rev. Geophys. Space Phys., 19*, 394-406, 1981.

Shelton, K.L., and P. Orville, Formation of synthetic fluid inclusions in natural quartz, *Am. Mineral., 65*, 1233-1236, 1980.

Sibson, R.H., Fluid flow accompanying faulting: field evidence and models, in *Earthquake Prediction: An International Review, Maurice Ewing Ser., Vol 4*, edited by D.W. Simpson and P.G. Richards, pp. 593-604, AGU, Washington, D. C., 1981.

Simmons, G., and D. Richter, Microcracks in rocks, in *The Physics and Chemistry of Minerals and Rocks*, edited by R.G.J. Strens, pp. 105-137, Wiley-Interscience, New York, 1976.

Smith, D.L., and B. Evans, Diffusional crack healing in quartz, *J. Geophys. Res., 89*, 4125-4135, 1984.

Sprunt, E.S. and A. Nur, Reduction of porosity by pressure solution: experimental verification, *Geology, 4*, 463-466, 1976.

Sprunt, E.S. and A. Nur, Destruction of porosity through pressure solution, *Geophysics, 42*, 726-41, 1977a.

Sprunt, E.S. and A. Nur, Experimental study of the effects of stress on solution rate, *J. Geophys. Res., 82*, 3013-22, 1977b.

Sprunt, E.S. and A. Nur, Microcracking and healing in granites: new evidence from cathodoluminescence, *Science, 205*, 495-497, 1979.

Tada, R. and R. Siever, Pressure solution during diagenesis, *Ann. Rev. Earth Planet. Sci. 17*, 89-118, 1989.

Thompson, A. H., Fractals in rock physics, *Annu. Rev. Earth Planet. Sci., 19*, 237-262, 1991.

Thirumalai K. and S.G. Demou, Thermal expansion behavior of intact and thermally fractured mine rocks, In: *Am. Inst. Phys. Conf. Proc.*, R.E. Taylor and G.L. Denmoan (Eds.), 60-71, 1973.

Tsang, Y. W., and P. A. Witherspoon, Hydromechanical behavior of a deformable rock fracture subject to normal stress, *J. Geophys. Res., 86*, 9287-9298, 1981.

Underwood, E.E., *Quantitative Stereology*, 274 pp., Addison Wesley, Reading, MA., 1970.

Vaughan, P.J., D.E. Moore, C.A. Morrow, and J.D. Byerlee, Role of cracks in progressive permeability reduction during flow of heated aqueous fluids through granite, *J. Geophys. Res., 91*, 7517-7530, 1986.

Walder, J. and Nur. A., Porosity reduction and crustal pore pressure development. *J. Geophys. Res. 89*, 11,539-11,548, 1984.

Walsh, J. B.,The effects of cracks on the compressibility of rock, *J. Geophys. Res. 70*, 381-389, 1965.

Walsh, J. B.,The effect of pore pressure and confining pressure on fracture permeability, *Int. J. Rock Mech. Min. Sci. Geomech. Abstr., 18*, 429-435, 1981.

Walsh, J. B., and M. A. Grosenbaugh, A new model for analytizing the effect of fractures on compressibility, *J. Geophys. Res., 84*, 3532-3536, 1979.

Walsh, J. B., and W. F. Brace, The effect of pressure on porosity and the transport properties of rock, *J. Geophys. Res., 89*, 9425-9431, 1984.

Walther, J.V. and P.M. Orville, Volatile production and transport in regional metamorphism, *Contrib. Mineral. Petrol., 79*, 252-257, 1982.

Wanamaker, B.J., and B. Evans, Experimental diffusional crack healing in olivine, in *Point Defects in Minerals, AGU Monogr. 31*, edited by R. N. Schock, pp. 194-210, AGU, Washington, D.C., 1985.

Wang, H.F., B. P. Bonner, S.R. Carlson, B. J. Kowallis, and H. C. Heard, Thermal stress cracking in granite, *J. Geophys. Res., 94*, 1745-1758, 1989.

Watson, E. B., J. M. Brenan, and D. R. Baker, Distribution of fluids in the continental lithosphere mantle: in *The Continental Lithospheric Mantle*, edited by, M. A. Menzies, pp. 111-125, Oxford University Press, 1991.

Watson, E. B. and Brenan, J. M., Fluids in the lithosphere, 1. Experimentally-determined wetting characteristics of $CO_2$-$H_2O$ fluids and their implications for fluid transport, host-rock physical properties, and fluid inclusion formation, *Earth Plan. Sci. Let. 85*, 497-515, 1987.

Williamson, E.D. and L.H. Adams, Temperature distribution in solids during heating and cooling, *Physical Review, 19*, 99-114, 1919.

Wong, T.-f., J.T. Fredrich and G.D. Gwanmesia, Crack Aperture Statistics and Pore Space Fractal Geometry of Westerly Granite and Rutland Quartzite: Implications for an Elastic Contact Model of Rock Compressibility, *J. Geophys. Res., 94*, 10267-10278, 1989.

Yong, C. and C-y. Wang, Thermally induced acoustic emission in Westerly Granite, *Geophys. Res.Lett.*, *7*, 1089-92, 1980.

Zhang, J., T.-f. Wong, and D. M. Davis, Micromechanics of pressure-induced grain crushing in porous rocks, *J. Geophys. Res.*, *95*, 341-352, 1990.

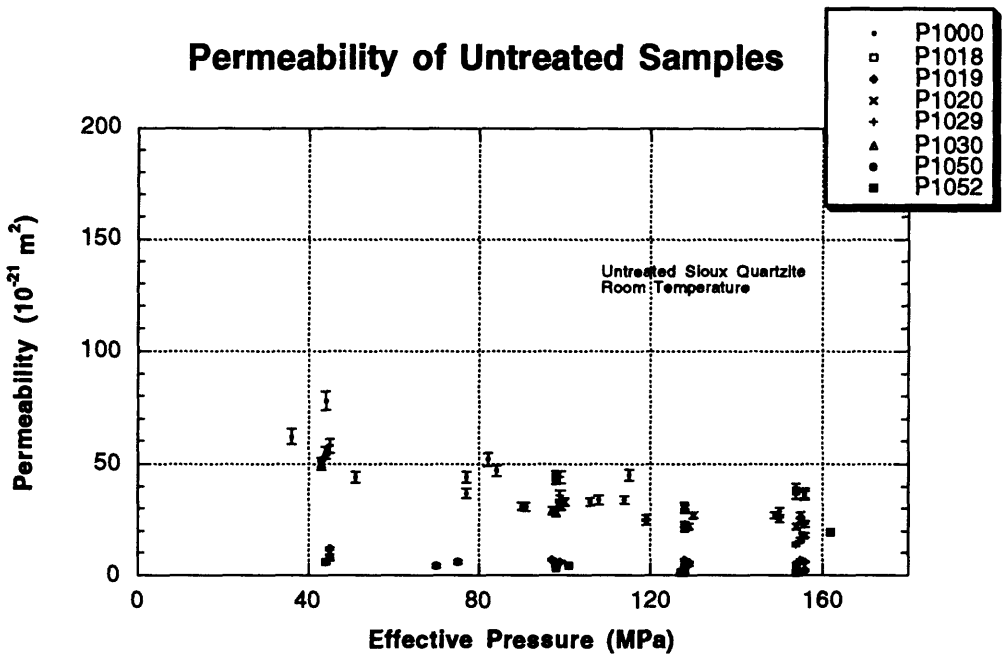



Figure 1. Permeability of untreated samples as a function of effective pressure.

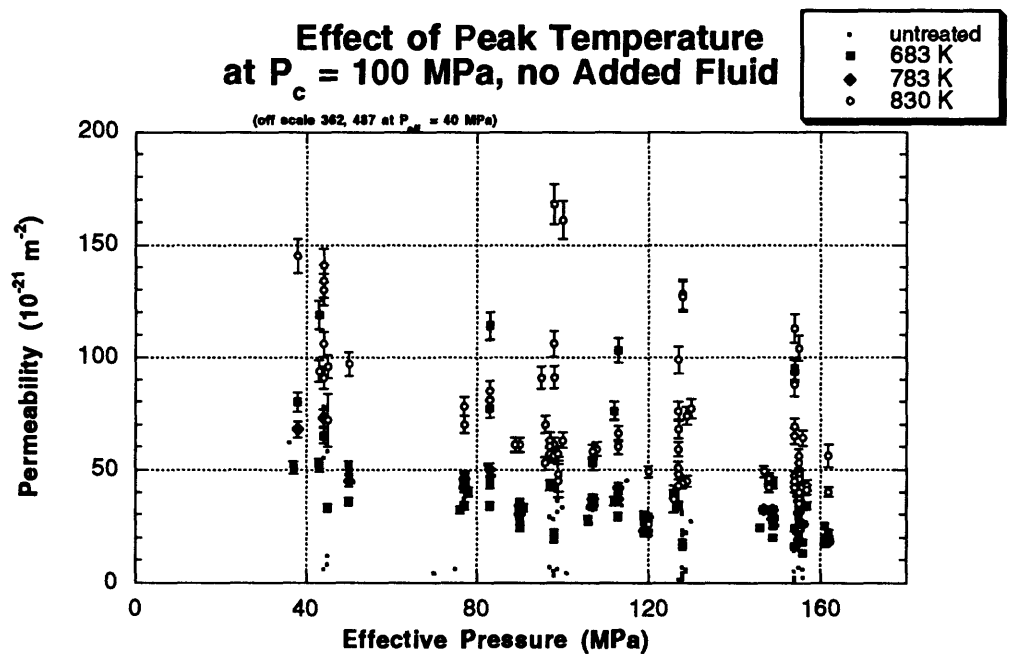



Figure 2. Permeability vs. effective pressure for samples heated to various temperatures at 100 MPa confining pressure with no pore fluid present. Note outliers with permeabilities of 362 and 487 nd (one sample) at $P_{eff}$ of 40 MPa.

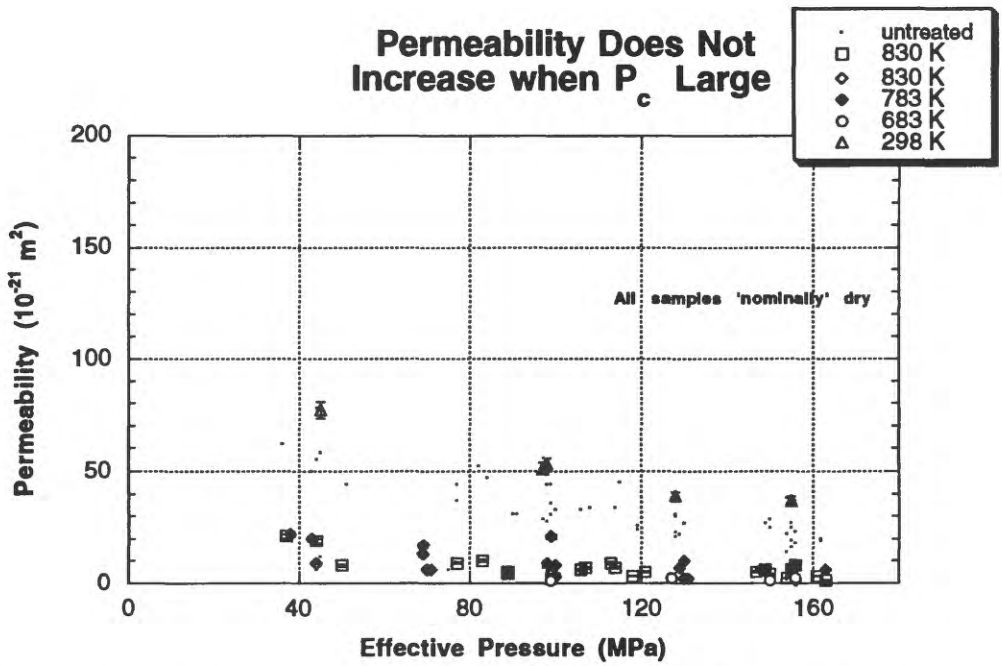



Figure 3. Permeability vs. effective pressure for samples heated to various temperatures at 200 MPa confining pressure.

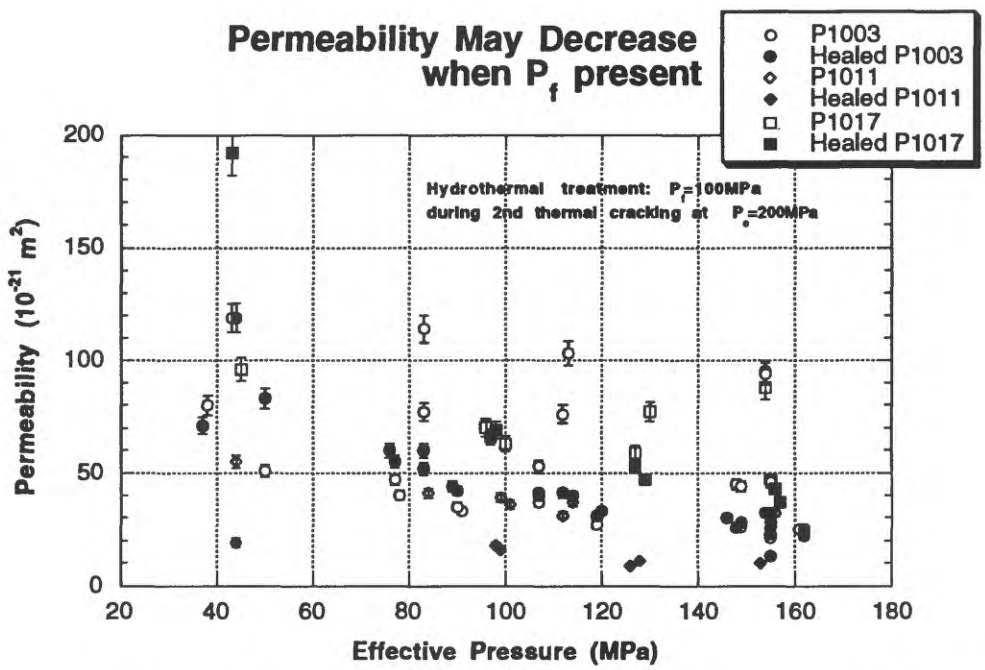



Figure 4. Permeabilities of samples heated to 830 K with and without pore fluids. The effective pressure in both cases was 100 MPa. For fluid saturated experiments, pore fluid pressure was 100 MPa.

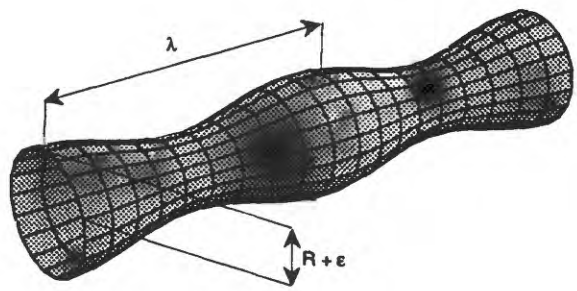



Figure 5a. A piece of infinite cylinder with sinusoidal perturbation which is characterized by wavelength and amplitude. Both parameter are normalized by the radius.

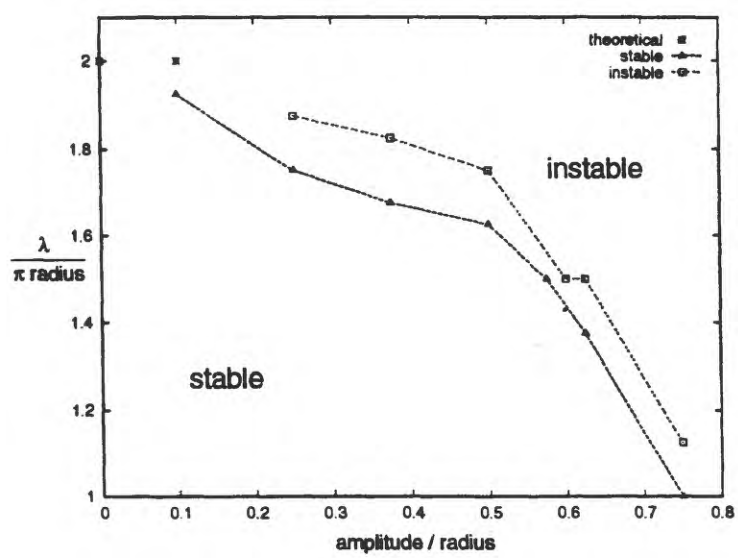



Figure 5b. Stability plot of an infinite cylinder with variations in radius along its length. The vertical axis shows the wavelength of the perturbation, the horizontal axis, the amplitude. The stars denote the stability criterion determined by *Nichols and Mullins* [1976a,b]. The triangles denote the last stable configuration tested and the squares indicate the first unstable configuration encountered. Somewhere in the region between these two zones the behavior changes from stable spheroidization (the tube evolves to one large sphere) to unstable ovulation (several fluid inclusions are formed)

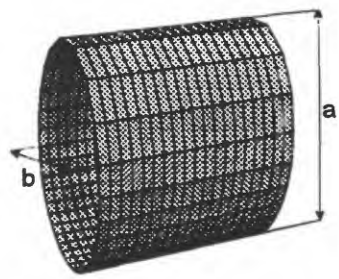



Figure 6a. A second family of shapes which were tested: ellipsoidal prisms. The ellipsoids are characterized by the major and minor semi-axies. The one shown here has axes of ratio 1.8 : 0.56. The eccentricity is 0.95.

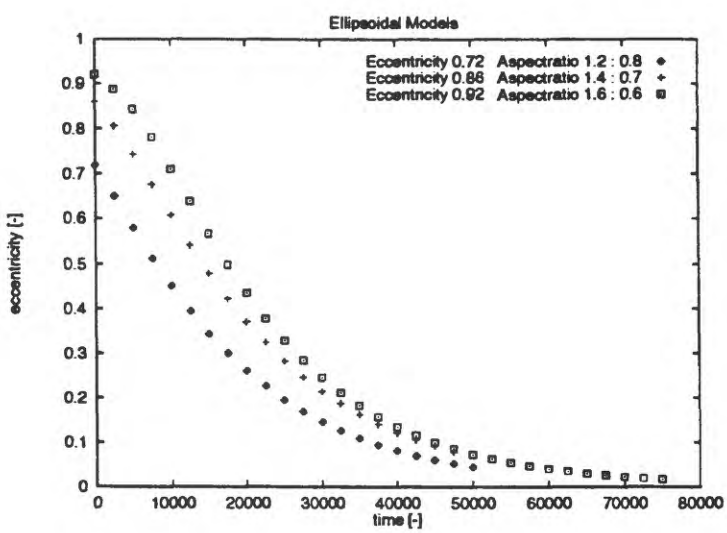



Figure 6b. The change of linear eccentricity with time for three ellipsoidal prisms of differing eccentricity. For small eccentricities the decay becomes exponential. Flatter shapes show instabilities. The last stable model had an eccentricity of ).92 or an axes ratio of 1.6 : 0.63. When the eccentricity was increased to 0.94, the crack healed unstabily forming a cylindrical tube at the edge of the ellipse.

# Fault Fluid Compositions from Fluid Inclusion Observations

**W. T. PARRY**

Department of Geology and Geophysics, University of Utah, Salt Lake City

## ABSTRACT

Hydrothermal alteration minerals and fluid inclusions have been examined in the exhumed footwall rocks of the Wasatch, Corral Canyon, and West Stansbury faults in Utah, the Teton fault, Wyoming, the Dixie Valley fault, Nevada, and the White Wolf fault, California. Hydrothermal minerals include early K-feldspar and biotite, followed by later epidote, chlorite, and muscovite, and latest prehnite, laumontite, stilbite, and clay.

Fault fluid compositions preserved in several types of fluid inclusions are extremely variable; a typical fault fluid does not exist. Components in the fluid include NaCl, $CO_2$, $CaCl_2$ in addition to $H_2O$. NaCl concentrations vary from 0 to 39 wt. %, $CaCl_2$ concentrations range up to 19 wt. % and $CO_2$ concentrations range up to 32 mole % in type II fluid inclusions, but some inclusions are present that are 100 mole % $CO_2$. Homogenization temperature measurements and pressure estimates confirm that these fluids were trapped at depth on the faults. Diverse fluid compositions are present with similar homogenization temperatures and estimated pressures suggesting similar depths on the faults.

The effects of fluids on faulting are related to the physical and chemical properties of the fluid. Pressure, temperature and fluid composition determine the solubilities of crack sealing minerals calcite and quartz and the formation of alteration minerals that are related to the mechanical behavior of the rock. Higher salinities increase solubilitites of calcite and quartz and decrease the pH for equilibrium among feldspars, muscovite and solution.

## INTRODUCTION

Hydrothermal fluid chemical interaction with fault zone rocks is evident in fault zones exhumed by erosion. Fractures are filled with precipitated minerals, alteration product minerals are present in the host rock, and secondary fluid inclusions are trapped in vein and matrix minerals. Hydrothermal fluids play a significant role in determining the mechanical stability of faults. Fault stability is affected by healing and sealing of cracks with precipitated minerals, by alteration product minerals with lower frictional or flow strength, by elevated fluid pressures where stability is governed by effective stress, and by fluid physical properties such as density and bulk modulus. The chemical composition, temperature, and pressure of the fluids determine chemical reactions with host rock, mineral saturations, and fluid density and bulk modulus. Characteristics of the fluid inclusions together with associated hydrothermal alteration product minerals constrain the fluid chemical composition, T, and P. We report here the chemical compositions of fluid inclusions associated with the Wasatch, West Stansbury, and Corral Canyon faults, Utah; the Dixie Valley fault, Nevada, the Teton fault, Wyoming, and the White Wolf fault, California.

Identity, density, and homogenization temperatures of fluid inclusion contents were

determined by observation of phase changes in doubly polished plates on a heating-freezing microscope stage using procedures outlined by Roedder (1984), Parry and Bruhn (1986), Parry (1986) and Parry et al. (1988, 1991). Phase changes that were used to determine fluid composition were the temperature of the $CO_2$ solid-liquid-vapor triple point, clathrate melting temperature; melting of ice, hydrohalite, and halite; $CO_2$ liquid-vapor homogenization; and overall homogenization of fluid inclusion contents.

Several types of fluid inclusions have been observed in fault systems: moderate salinity inclusions (type I) with salinities from 0 to 20 equivalent weight percent NaCl, carbon dioxide bearing inclusions (type II) that contain up to 30 mole percent $CO_2$ in addition to a saline aqueous fluid, low eutectic temperature inclusions (type III) with initial melting temperatures as low as 45°C suggesting a significant $CaCl_2$ component in the fluid, and inclusions containing a daughter halite crystal (type IV) in addition to a vapor bubble.

Salinities of type I fluid inclusions were estimated from ice melting temperatures and the regression equation of Potter et al. (1978). Salinities of type II inclusions were estimated from clathrate melting temperatures (Collins, 1979; Bozzo et al., 1975). Salinities of type III inclusions were estimated using the equations and coefficients of Oakes et al. (1990), and salinities of type IV inclusions on the halite liquidus were calculated using the equation and coefficients of Sterner et al. (1988). Densities and bulk moduli of fluid inclusion fluids at their homogenization temperatures were estimated using equations and data of Potter and Brown (1977) for type I fluids and the MRK equation of state and data presented by Bowers and Helgeson (1983) for type II fluids. Solubility of quartz as a function of fluid salinity and temperature was calculated using the equations of Fournier (1983), and estimated pH of fluids in equilibrium with quartz, K-feldspar, albite, and muscovite was calculated using equilibrium constants calculated with SUPCRT92 (Johnson, Oelkers, and Helgeson, 1991) and activity coefficients calculated using equation 298 in Helgeson et al., 1981).

The greatest source of error in fluid inclusion measurements is a consequence of post entrapment changes. Dissolution and reprecipitation of host mineral surrounding the trapped fluid result in a necked inclusion. If the necking takes place after separation of a vapor phase, then the result is varying liquid to vapor ratios and varying density of fluid inclusion contents. Necking would not alter the salinity of the liquid phase. Measurements of fluid inclusions with liquid vapor ratios affected by necking have been avoided by avoiding fluid inclusions in close proximity to one another with widely varying liquid to vapor ratios and by checking nearby inclusions with apparently similar liquid- vapor ratios to insure similar homogenization temperatures.

Establishing that the fluid inclusions represent fault fluids has been accomplished in several ways. Solid inclusions of calcite, chlorite, and sericite occasionally accompany the fluid inclusions, and sericite veinlets in surrounding feldspars are often coextensive with fluid inclusion planes. Fluid inclusion planes in quartz bisect the angle between the c-axis and deformation lamellae suggesting the fluid inclusion planes parallel the direction of maximum principal stress. The orientations of fluid inclusion trails measured on oriented samples are consistent with the fault stress system. Fluid inclusion abundances vary systematically with proximity to the fault, for example quartz grains nearest the Dixie Valley fault contain up to $2x10^5$ type I fluid inclusions per cubic millimeter and abundance decreases to $3x10^4$ fluid inclusions per cubic millimeter within 6 m of the fault. Type I and type II fluid inclusions were sufficiently abundant to count on the Teton fault; in addition fluid inclusions consisting of only a single liquid phase could be counted. Type II fluid inclusions were present only in samples

within 300 m of the fault in the Broken Falls area. Here abundance of type II fluid inclusions is 144 to 663 fluid inclusions per cubic millimeter of quartz. Abundance of type I inclusions in these samples is 1400 to 2800 inclusions per cubic millimeter. A significant limitation in establishing the relationship between fluid inclusions and faulting is our failure to establish a fluid inclusion chronology. We are unable to determine the relative ages of the various type of fluid inclusions in a given sample.

## The Wasatch Fault, Utah

Fluid inclusions and alteration minerals have been examined in the southern portion of the Salt Lake segment of the Wasatch Fault in an Oligocene quartz monzonite (Parry and Bruhn, 1986; Parry et al., 1988). Vein filling and pervasive alteration minerals are biotite, K-feldspar, chlorite, epidote, muscovite, laumontite, prehnite, calcite, and clay minerals. Secondary fluid inclusions are present in healed microfractures in igneous quartz. Type II fluid inclusions associated with the chlorite and epidote alteration for which a full set of measurements have been obtained have salinities from 4.5 to 17.3 wt % NaCl equivalent and contain 3 to 32 mole percent $CO_2$ shown in Figure 1. A few fluid inclusions not plotted in Figure 1 contain up to 100 mole % $CO_2$. Homogenization temperatures of type II fluid inclusions range from 176°C to 353°C (Figure 2). Minimum fluid pressures (Figure 3) vary from 600 to 2950 bars. Fluid temperature, pressure, and composition evolved along a path from 350°C to 150°C, from lithostatic to hydrostatic pressure, and from more than 30 mole % $CO_2$ to less than 3 mole % $CO_2$ with continued displacement of the fault. Type I fluid inclusions are associated with laumontite, prehnite, and clay; type II fluid inclusions are associated with chlorite and epidote alteration. Salinities of type I fluid inclusions shown in Figures 4 range from 2 to 16 wt% NaCl equivalent and no $CO_2$ could be detected.

## The Dixie Valley Fault, Nevada

Fluid inclusions and alteration minerals have been examined in a portion of the footwall of the Dixie Valley Fault, Nevada where footwall rocks are Oligocene granites to granodiorites (Parry et al., 1991). Temporally and spatially overlapping hydrothermal alteration mineral assemblages occur as a narrow band near the 1954 rupture in the fault footwall. Early hydrothermal biotite and K-feldspar are followed by later Fe-chlorite and epidote then sericite and calcite. The latest hydrothermal minerals are stilbite, laumontite, kaolinite, alunite, smectite, illite and pervasive replacement of rock with fine grained quartz, chalcedony, and opal.

Secondary fluid inclusions are trapped in microfractures in igneous quartz. Type I, type II, type III, and type IV fluid inclusions are present. Type I fluid inclusion salinity ranges from 0.1 to 16.9 wt%NaCl (Figure 5), type II salinities range from 0.62 to 6.29 wt. % NaCl. Ice melting temperatures in type III fluid inclusions are -10.1°C to -26.0°C and initial melting temperatures are as low as -45°C and ice crystal nucleation temperatures are as low as -60°C to -80°C. These characteristics suggest the presence of a significant $CaCl_2$ component. Type III salinities are 12.9 to 25.3 wt. % NaCl+$CaCl_2$ with 30 to 55% of the total being NaCl. Type IV salinities are 30.1 to 39.2 wt % NaCl. Type II fluid inclusions for which pressure estimates could be made contain 3 to 16.7 mole % $CO_2$ (Figure 1) and were trapped at minimum pressures up to 1570 bars at minimum temperature of up to 340°C (Figures 1-3). A few inclusions not

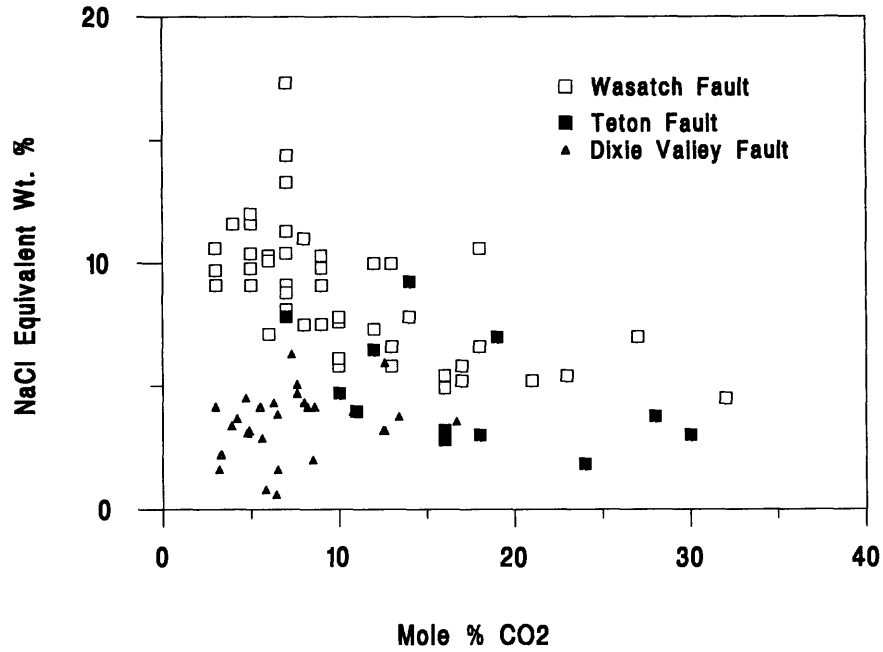



Figure 1. Mole fraction $CO_2$ and salinity of type II fluid inclusions from the Wasatch, Teton and Dixie Valley faults.

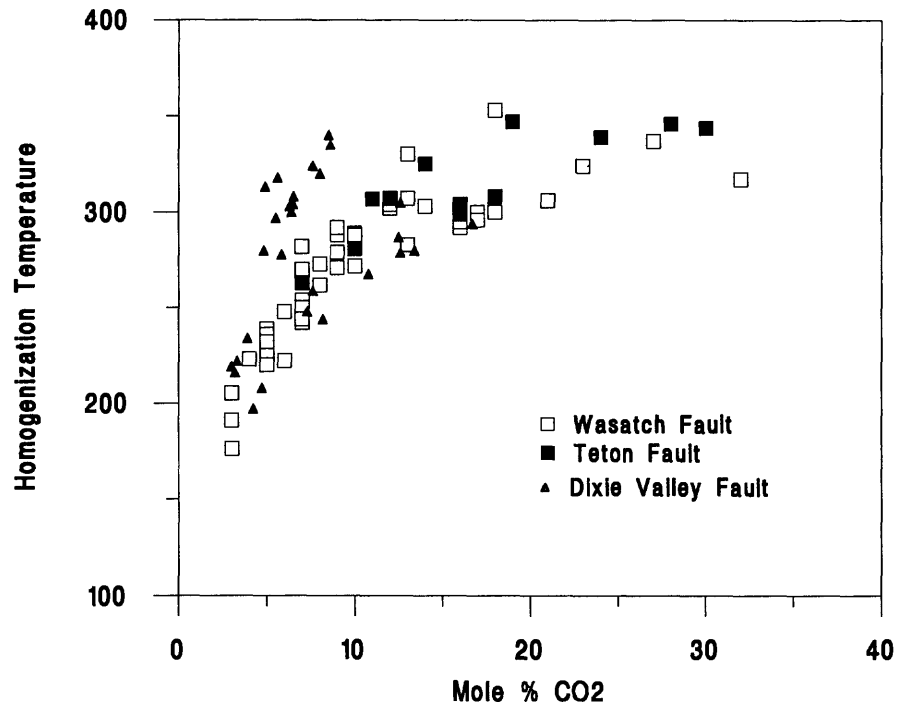



Figure 2. Temperature of homogenization and $CO_2$ content of type II fluid inclusions from theWasatch, Dixie Valley, and Teton faults.

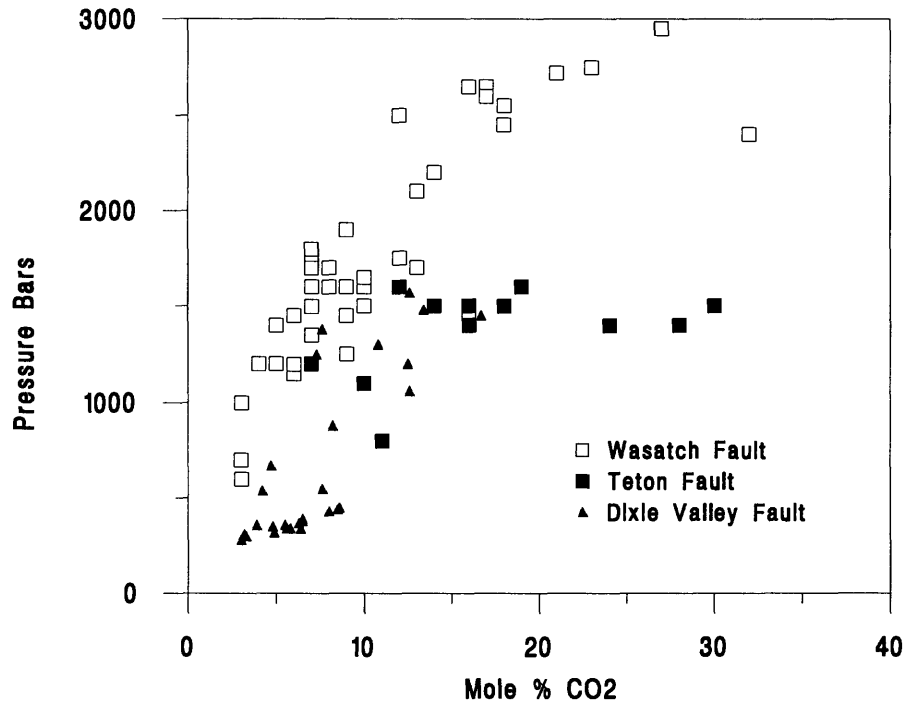



Figure 3. Minimum entrapment pressure and $CO_2$ content of type II fluid inclusions from the Wasatch, Dixie Valley, and Teton faults.

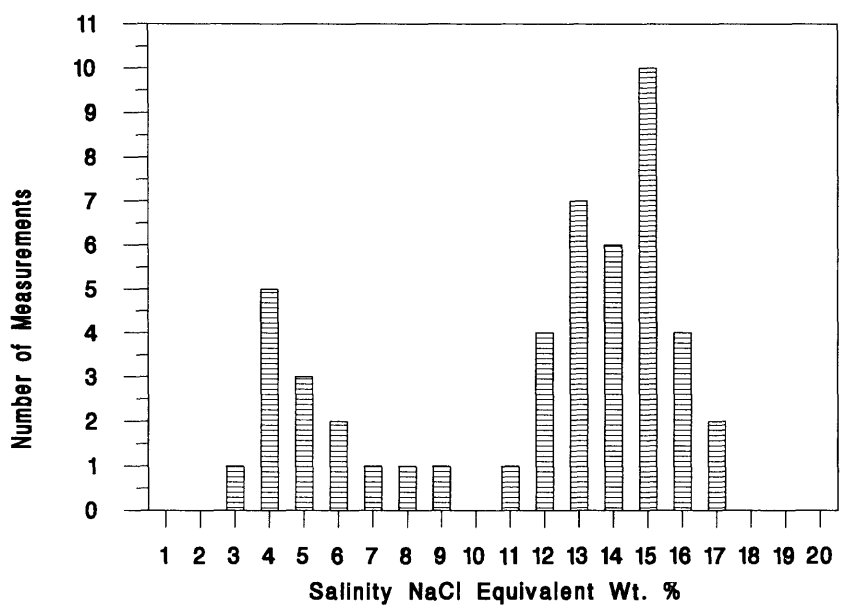



Figure 4. Salinity of type I fluid inclusions from the laumontite-prehnite alteration assemblage on the Wasatch fault, Utah.

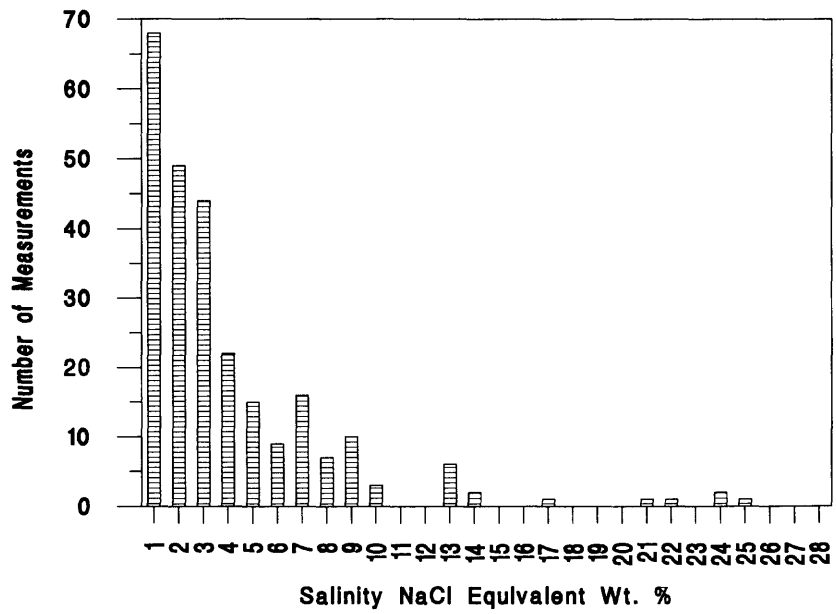



Figure 5. Salinity of type I and type III fluid inclusions from the Dixie Valley fault, Nevada.

plotted on the figures contain up to 51 mole % $CO_2$.

## THE TETON FAULT, WYOMING

Fluid inclusions have been examined in the Precambrian age Mount Owens quartz monzonite footwall rocks of the Teton fault exposed in the central Teton Range in northwestern Footwall rocks of the Teton fault in the central portion of the Teton Range consist of (in order of increasing age) dikes of tholeiitic diabase and basalt, the Mount Owen quartz monzonite, the Webb Canyon gneiss, Rendezvous Metagabbro, and Layered gneiss (Reed and Zartman, 1973; Embree, 1976). The Mount Owen quartz monzonite, exposed in the central Teton Range from Leigh Canyon on the north to Deaths Canyon on the south consists of 20 to 30% microcline clouded with sericite, 25 to 35% plagioclase (An30-40), 30 to 40 % strained quartz, 1 to 10 % biotite and up to 5% muscovite. The rock is non foliated and displays cataclastic and crystalloblastic textures. The layered gneiss, Webb Canyon gneiss and Rendezvous metagabbro all record a major metamorphic event at 2875±150Ma dated by Rb-Sr methods. K-Ar age of hornblende is 2780 and 2800 Ma. The Mount Owen quartz monzonite crystallized at 2495±75 Ma (Rb-Sr). Partial equilibration of the Rb-Sr system occurred at 1800 Ma. K-Ar ages of biotite and muscovite throughout the Teton Range demonstrate the presence of a wide spread thermal event at 1330 to 1500 Ma (Reed and Zartman, 1973; Embree, 1976). Faulting in the Precambrian is concurrent or post dates this 1330-1500 Ma event.

Samples of Mount Owen quartz monzonite display the type of alteration observed elsewhere in the Teton Range. Muscovite occurs as an alteration product mineral as fine grains that are disseminated throughout the feldspars or as coarser grains in veinlets. Epidote and chlorite commonly replace biotite and hornblende. Prehnite rarely occurs as a replacement of biotite crystals. Fluid inclusions are present in quartz in the quartz monzonite and in quartz veins as planar arrays of secondary fluid inclusions. Type II fluid inclusions contain 1.83 to 9.24 wt % NaCl and 7 to 30 mole % $CO_2$ (Figure 1). Homogenization temperatures are 263°C to 347°C, and estimated minimum entrapment pressures are 800 to 1600 bars (Figure 2 and 3). Type III fluid inclusions fluids contain 12 to 24 NaCl equivalent wt. %, but the dissolved salt is dominantly $CaCl_2$ (20% NaCl, 80% $CaCl_2$).

## THE WHITE WOLF FAULT, CALIFORNIA

The White Wolf fault separates the steep northwest slope of the Tehachapi Mountains from the southern San Joaquin Valley southeast of Bakersfield, California. Uplift and erosion of the hangingwall of the White Wolf fault has exhumed granitic rock from depth that contains evidence of episodic fluid flow. Preliminary work on exhumed fault rock (Hitchcock, 1993) has identified extensive hydrothermal alteration, abundant quartz and calcite mineral veins, and secondary fluid inclusions that provide a record of fluid migration along the fault. Alteration minerals in exhumed granitic rock include K-feldspar, chlorite, sericite, laumontite, prehnite, and calcite. Numerous quartz veins were examined for suitable fluid inclusions for measurement, but most of the quartz is so highly strained that no fluid inclusions remain. Secondary fluid inclusions were found in quartz in two samples of altered granitic rocks. These fluid inclusions belong to Type I with salinities that range from 2.2 to 6 NaCl equivalent wt. %. Homogenization temperatures are 137°C to 272°C shown in Figure 6. Fluid inclusions showed no evidence of

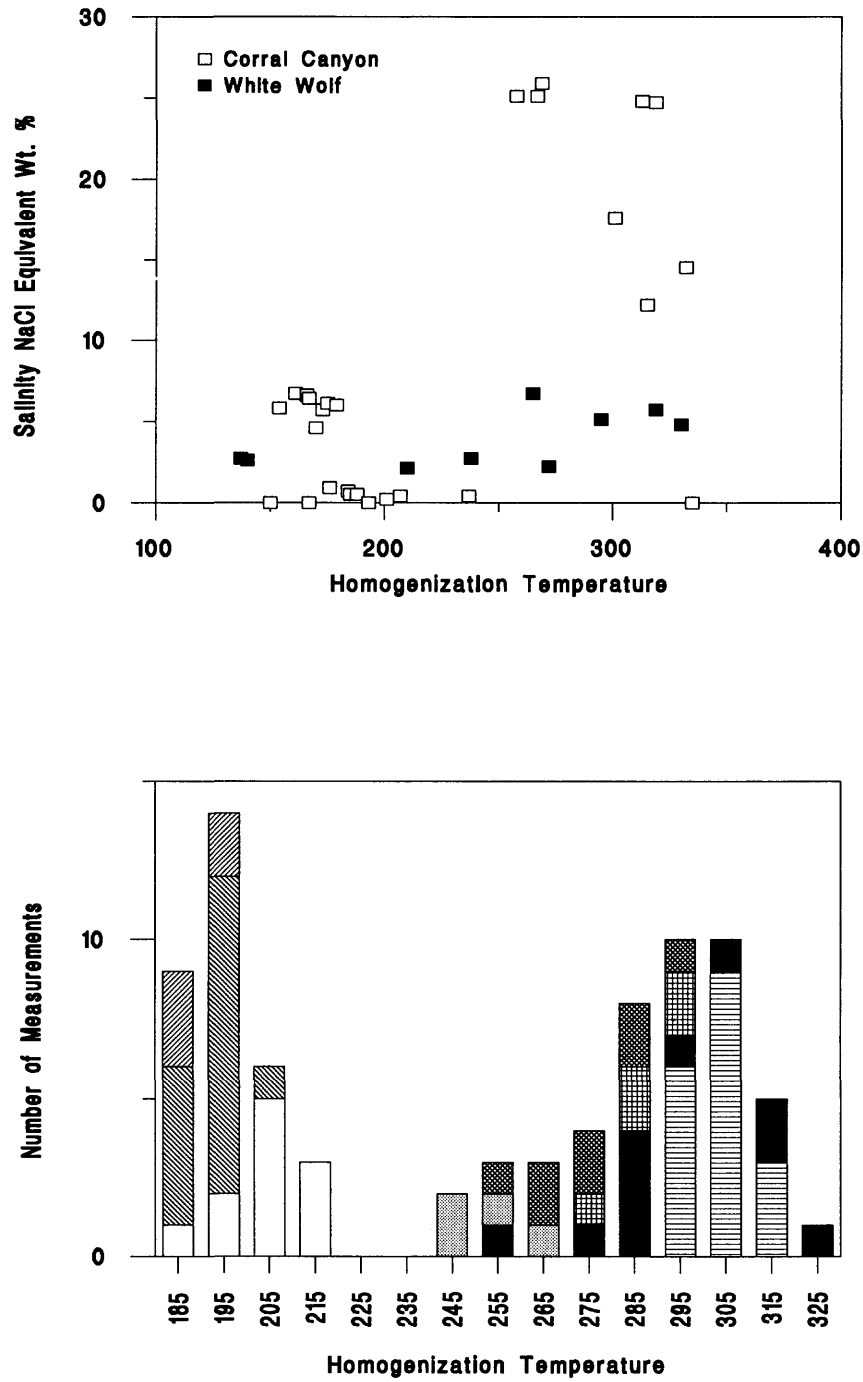



Figure 6. Top: Salinity and homogenization temperatures of type I fluid inclusions from the White Wolf fault, California and the Corral Canyon shear zone, Utah. Bottom: Histograms of homogenization temperatures of type IV fluid inclusions from the Corral Canyon shear zone. Each fill pattern represents fluid inclusions from a separate healed fracture.

$CO_2$, but abundant calcite in the fault zone demonstrates the presence of $CO_2$ in the fault fluids.

## THE CORRAL CANYON SHEAR ZONE, UTAH

The Corral Canyon shear zone is one of several high-level denudation faults resulting from extension in the eastern Great Basin that cuts Tertiary granite on the west flank of the Mineral Mountains in southwestern Utah. Alteration and studies of type IV fluid inclusions are reported by Cady (1982). In less deformed but highly sheared granite, sericite, chlorite, and epidote replace plagioclase and biotite. Greater abundances of sericite, chlorite and epidote with calcite are observed in more intensely deformed cataclasite. Fluid inclusions occur in vein quartz or in healed microfractures in primary igneous quartz. The veins and microfractures are aligned with the major planes of failure within the shear zone. Salinities of Type I fluid inclusions range from 0 to 25.8 wt. % NaCl equivalent and homogenization temperatures are 150°C to 335°C. Type IV fluid inclusions contained a small halite crystal that was too small to observe melting behavior in the heating stage. Salinities of these inclusions calculated from an estimate of the size of the halite crystal were from 26.6 to 27.8 %NaCl. The inclusions homogenized by vapor disappearance at temperatures of from 155°C to 335°C (Figure 6). Several fluid inclusions were measured by Cady (1982) on each of 8 different healed fractures with the results shown in Figure 6. Fluid inclusions on some fractures consistently homogenized at temperatures near 200°C and others homogenized at temperatures as high as 300°C. Homogenization temperatures of fluids with very low salinities also homogenized at temperatures from 150°C to nearly 350°C. We were not able to establish a fluid inclusion chronology that would indicate the relative ages of low and high salinity fluids.

## THE WEST STANSBURY FAULT, UTAH

The west Stansbury Fault forms the boundary between Skull Valley and the Stansbury Mountains in north central Utah. A zone of scarps is nearly continuous for 30 km along the north striking fault zone (Barnhard, 1988). Footwall rocks contain abundant quartz and calcite veins associated with faulting that are host for secondary fluid inclusions. Histograms of fluid salinities are shown in Figure 7. Fluid inclusions in calcite have relatively low salinities from 0 to 3 wt % NaCl and homogenization temperatures from 120°C to 340 °C. Fluid inclusions in vein quartz have higher salinities from 7 to 18 wt % NaCl and homogenization temperatures from 100°C to 230°C. The homogenization temperatures for fluid inclusions in calcite and quartz overlap; the salinities do not.

## DISCUSSION

Fluid inclusions and hydrothermal alteration minerals record the presence of aqueous fluids at depth in a variety of host rocks and faults. Fluid composition is extremely variable from near fresh water to high salinities of 39.2 wt. % NaCl and include $CaCl_2$ and $CO_2$. The source of dissolved constituents is not known, but the present day Dixie Valley fault fluid involves evaporites from the nearby basin. A typical fault fluid does not exist. Fluid variability is as great within footwall rocks of an individual fault as among all of the faults studied. Also, fluid compositions vary at similar minimum temperatures and pressures on individual faults.

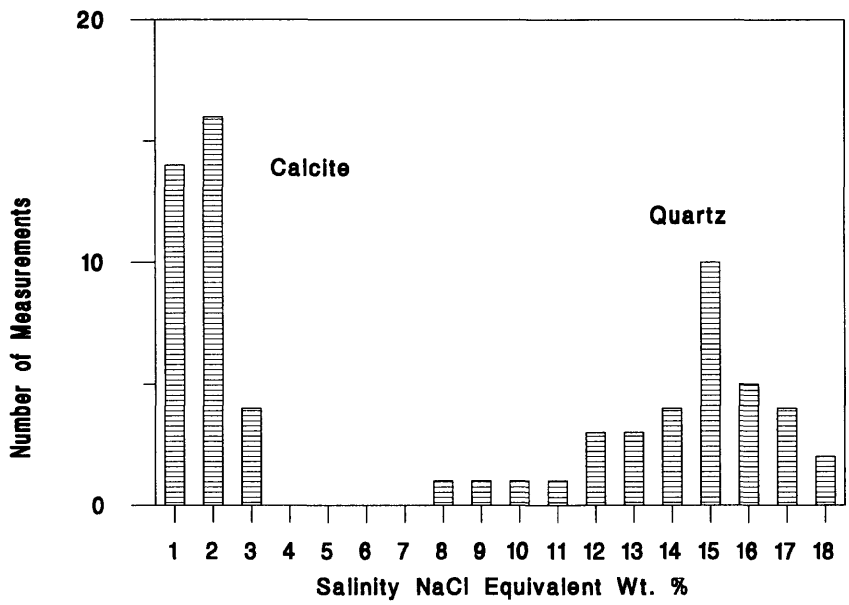



Figure 7. Salinity of type I fluid inclusions from quartz and calcite from the West Stansbury fault, Utah

The physical effects of fluids on faulting are related to fluid physical properties density and compressibility that are determined by fluid composition in addition to T and P. The chemical effects of fluids on faulting are also related to fluid chemical composition. Dissolved $CO_2$ produces ionizable hydrogen that drives many chemical reactions such as the conversion of feldspar to mica.

Densities calculated from homogenization temperature and salinity of fluid inclusions are shown on Figure 8 together with type I fluid inclusions from the Wasatch fault, Utah. Densities of fluids vary considerably. At temperatures near 300°C on the Wasatch fault, fluid densities from 0.75 g/cm$^3$ to 0.9 g/cm$^3$ are shown. Such variability in fluid density should play a significant role in fluid flow patterns. The bulk modulus of fault fluids at temperatures near 300°C calculated by Parry and Bruhn (1990) is $0.8 \times 10^3$ MPa to $1.91 \times 10^3$ MPa depending on fluid composition and pressure. Separation of $CO_2$ or steam during pressure reduction can result in a drastic reduction in bulk modulus. For example decompression of a type II fluid on the Dixie Valley fault from 145 MPa to 37 MPa results in separation of $CO_2$ and a bulk modulus of the two phase fluid that is 90 % lower than the original single phase fluid.

Solubilities of quartz and calcite are related to the fluid salinity. Dissolved NaCl greatly increases the solubility of quartz above about 300°C (Fournier,1985a) affecting the sealing of fractures by precipitated quartz. Solubility of quartz as a function of salinity and temperature were calculated using the equations and data of Fournier (1983) for pressures on the liquid-vapor curve (boiling) and are shown in Figure 8. Salinity plays a role in quartz solubility in the temperature range where the concentration isopleths have negative slopes mostly above about 300°C, for example, quartz solubility decreases from 1000 mg/l to 800 mg/l at 350°C for a change in salinity from 10 wt. % to 1 wt. %. A temperature change of approximately 40°C results in a similar change in solubility at 350°C. The primary control of quartz solubility in fault zones is temperature, but mixing quartz-saturated fluids of contrasting salinity could result in quartz precipitation and fracture sealing.

The solubility of calcite in fault fluids is determined by temperature, pH, salts in solution, and partial pressure of $CO_2$. At a constant $CO_2$ pressure, decreasing temperature will increase the solubility of calcite, decreasing $CO_2$ partial pressure will decrease calcite solubility, and adding NaCl to solutions will increase the solubility of calcite through the effect of decreasing the activity coefficients of aqueous species (Fournier, 1985b). For example, at 300°C pure water will dissolve about 0.1 millimole of calcite at 12 atmospheres $CO_2$ pressure, but the solubility increases to 1.6 millimoles of calcite in 5.5 wt. % NaCl (1 m) (Fournier, 1985b) In a fault system in which fluids cool and loose $CO_2$ during decompression, the loss of $CO_2$ predominates over temperature and calcite precipitates. Calcite in veins that seal fractures could have precipitated by several mechanisms. A rock fracturing event would open new porosity, lower fluid pressure, and cause effervescence of dissolved $CO_2$. The $CO_2$ loss would result in a decrease in calcite solubility and resultant precipitation of calcite. Calcite could also precipitate from cool solutions that are circulating into warmer regions of the fault, or dilution of a higher salinity fluid that is calcite saturated by a low salinity fluid.

The pH of solutions in equilibrium with albite, K-feldspar, and muscovite is a function of solution salinity, because of the effect of salinity on activity coefficients of aqueous species in the reaction (Henley and Brown, 1985). The pH of these solutions were calculated using SUPCRT92 for computation of the equilibrium constants for reactions 1 and 2 below and the activity-concentration model of Helgeson et al. (1981). The calculations are illustrated in Figure 9 in the temperature range 100°C to 400°C at 1 kb for salinities in terms of NaCl equivalent

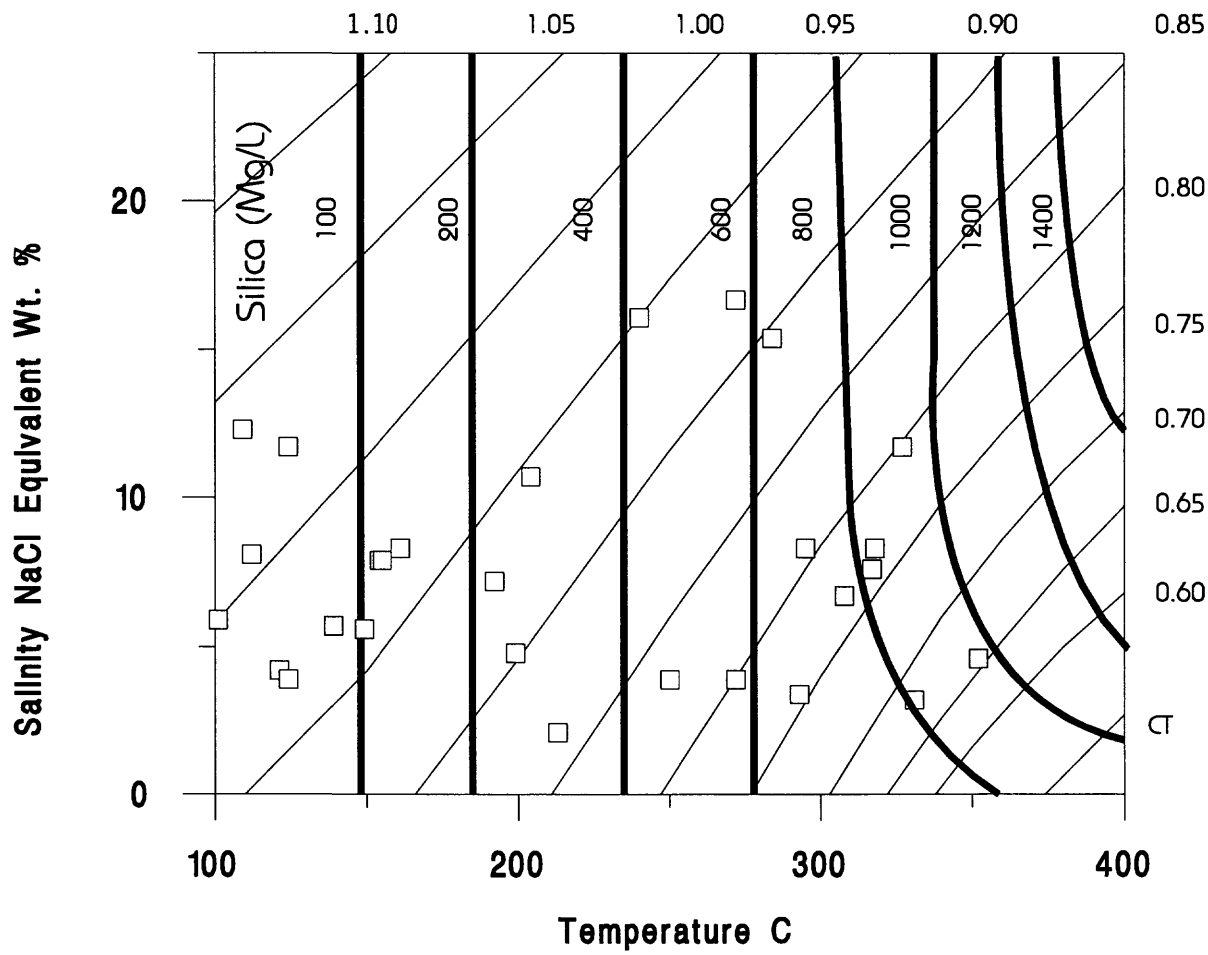



Figure 8. Solubility of quartz as a function of temperature and salinity (heavy lines) and densities in gm/cm$^3$ of fault fluids (light lines). Open squares are salinity and homogenization temperature of type I fluid inclusions on the Wasatch fault.

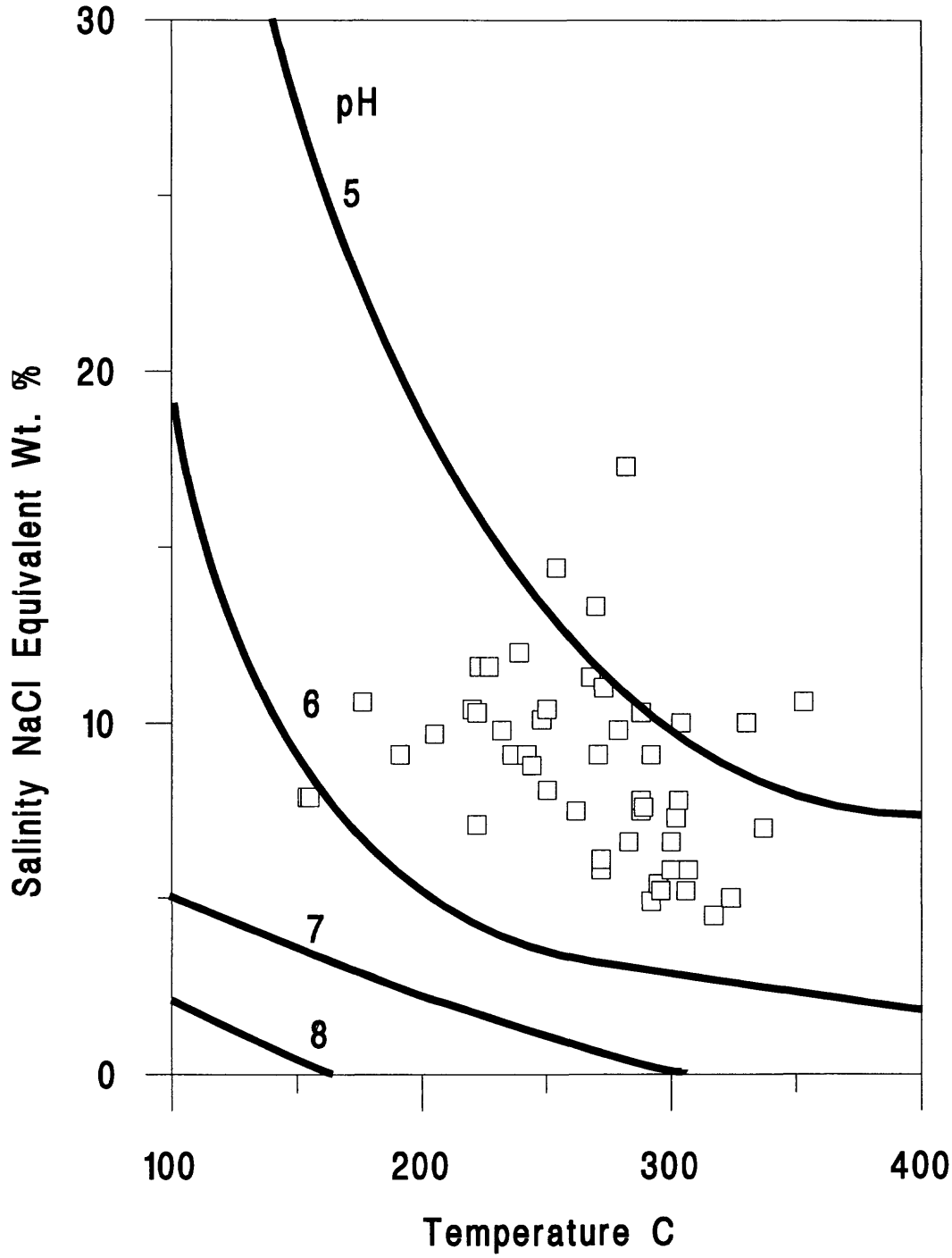



Figure 9. Calculated pH as a function of fluid salinity for equilibrium among albite, K-feldspar, and muscovite. Salinity and homogenization temperature of type II fluid inclusions from the Wasatch fault, Utah are plotted as open squares.

weight percent.

$$3KAlSi_3O_8 + 2H^+ = KAl_3Si_3O10(OH)_2 + 6SiO_2 + 2K^+ \quad (1)$$
$$KAlSi_3O_8 + Na^+ = NaAlSi_3O8 + K^+ \quad (2)$$

The figure shows that pH is lower for higher salinities and higher temperatures. Fluid inclusion temperatures and salinities for type II fluid inclusions on the Wasatch fault are shown for comparison. These fluids would lie in the pH range 5 to 5.5 in equilibrium with feldspars and muscovite.

Fault fluids on the Wasatch, Dixie Valley, and Teton faults appear to evolve in composition from high to low $CO_2$ contents with continued upward displacement of the footwall and decrease in pressure and temperature (Parry et al., 1991, Parry et al., 1988, Parry and Bruhn, 1990). The fluids represented by fluid inclusions play a key role in the mechanical behavior of faults. Effective stress is reduced by pore fluid pressure that exceeds hydrostatic pressure, in fact rupture may be initiated by high fluid pressure. Newly formed fractures reduce the fluid pressure, CO2 may effervesce, and hydrothermal minerals precipitate sealing the fractures and reducing rock permeability.

## References


Barnhard, T. P., Fault-scarp studies of the Stansbury Mountains, Utah:In the Footsteps of G. K. Gilbert--Lake Bonneville and neotectonics of the eastern Basin and Range Province, Utah Geological and Mineral Survey Miscellaneous Publication 88-1, 58-59, 1988.

Bowers, T. S., and H. C. Helgeson, Calculation of the thermodynamic and geochemical consequences of non ideal mixing in the system $H_2O$-$CO_2$-NaCl on phase relations in geologic systems: Equation of state for $H_2O$-$CO_2$-NaCl fluids at high pressures and temperatures, Geochim. Cosmochim. Acta, 47, 1247-1275, 1983.

Bozzo, A. T., H-S Chen, J. R. Kass, and A. J. Barduhn, The properties of the hydrates of chlorine and carbon dioxide, Desalination, 16, 303-320, 1975.

Cady, C. C., Hydrothermal alteration of the Corral Canyon shear zone, Mineral Mountains, Utah with inferences for shearing, M. S. thesis, University of Utah, Salt Lake City, 1982..

Collins,P. L. F., Gas hydrates in $CO_2$-bearing fluid inclusions and the use of freezing for estimation of salinity, Econ. Geol., 74, 1435-1444, 1979.

Embree, G. F., Structural analysis of the Precambrian rocks, Green Lakes Basin, Teton Range, Wyoming: Ph. D. Dissertation, University of Idaho, Moscow, Idaho, 104 pp., 1976.

Fournier, R. O., A method of calculating quartz solubilities in aqueous sodium chloride solutions: Geochim. et Cosmochim. Acta, 47, 579-586, 1983,.

Fournier, R. O., The behavior of silica in hydrothermal solutions: in Berger, B. R., and P. M. Bethke eds. Geology and Geochemistry of Epithermal Systems, Reviews in Economic Geology Volume 2, Society of Economic Geologists El Paso, Texas, 45-61, 1985a.

Fournier, R. O., Carbonate transport and deposition in the epithermal environment:in Berger, B. R., and P. M. Bethke eds. Geology and Geochemistry of Epithermal Systems, Reviews in Economic Geology Volume 2, Society of Economic Geologists El Paso, Texas, 63-72, 1985b.

Helgeson, H. C., D. H. Kirkham, and G. C. Flowers, Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures: IV. Calculation of activity coefficients, osmotic coefficients, and apparent molal and standard and relative partial

molal properties to 600°C and 5kb: American Journal of Science, 281, 1249-1516, 1981.
Henley, R., W., and K. L. Brown, A practical guide to the thermodynamics of geothermal fluids and hydrothermal ore deposits::in Berger, B. R., and P. M. Bethke eds. Geology and Geochemistry of Epithermal Systems, Reviews in Economic Geology Volume 2, Society of Economic Geologists El Paso, Texas, 25-44, 1985.
Hitchcock, C. S., Hydrologic controls on fluid flow following the 1952 Kern County earthquake, Kern County, California, M. S. thesis, University of Utah, Salt Lake City, 166p, 1993.
Johnson, J. W., E. H. Oelkers, and H. C. Helgeson, SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions among them as functions of temperature and pressure: Computers Geosci. 18, 899-947, 1991.
Oakes, C. S., R. J. Bodnar, and J. M. Simonson, The system NaCl-CaCl2-H2O, I, the ice liquidus at 1 atm total pressure, Geochim. et Cosmochim. Acta, 54, 603-610, 1990.
Parry, W. T., Estimation of $XCO_2$, P, and fluid inclusion volume from fluid inclusion temperature measurements in the system NaCl-$CO_2$-$H_2O$, Econ. Geol., 81, 1009-1013, 1986.
Parry, W. T., and R. L. Bruhn, Pore fluid and seismogenic characteristics of fault rock at depth on the Wasatch fault, Utah, J. Geophys. Res., 91, 730-744, 1986.
Parry, W. T., and R. L. Bruhn, Fluid pressure transients on seismogenic normal faults, Tectonophysics, 179, 335-344, 1990.
Parry, W. T., P. N. Wilson and R. L. Bruhn, Pore-fluid chemistry and chemical reactions on the Wasatch normal fault, Utah, Geochim. et Cosmochim. Acta, 52. 2053-2063, 1988.
Parry, W. T., D. Hedderly-Smith, and R. L. Bruhn, Fluid inclusions and hydrothermal alteration on the Dixie Valley Fault, Nevada: J. Geophys. Res., 96, 19733-19748, 1991.
Potter, R. W., II, and D. L. Brown, The volumetric properties of aqueous sodium chloride solutions from 0°C to 500°C at pressures up to 2000 bars based on regression of available data in the literature, U. S. Geol. Surv. Bull., 1421-C, 36 pp., 1977.
Potter, R. W., II, M. A. Clynne, and D. L. Brown, Freezing point depression of aqueous sodium chloride solutions, Econ. Geol., 73, 284-285, 1978.
Reed, J. C., Jr. and Zartman, R. E., Geochronology of Precambrian rocks of the Teton Range, Wyoming: Geological Society of America Bulletin v. 84, 561-582, 1973.
Roedder, E. Fluid inclusions, Reviews in Mineralogy, v. 12, 644 pages, Mineralogical Society of America, 1980.
Sterner, S. M., D. L. Hall, and R. J. Bodnar, Synthetic fluid inclusions, V, Solubility relations in the system NaCl-KCl-H2O under vapor saturated conditions, Geochim., Cosmochim. Acta, 52, 989-1005, 1988.

# The Geochemistry and Role of Fluids in Large Continental Structures: An Overview

Robert Kerrich and T. Kurtis Kyser
Department of Geological Sciences, University of Saskatchewan,
Saskatoon, Canada S7N 0W0

## INTRODUCTION

The interaction of fluids with the lithosphere governs a wide variety of geological phenomena. Advection of fluids through crustal rocks is a key determinant of heat and mass transport. Crustal rheology is highly sensitive to the presence of 'water' at the intracrystalline level, and fluids influence the rate of grain boundary diffusion, and the magnitude of effective stress at the intercrystalline level (Fyfe et al., 1978; Etheridge et al., 1983). Seismic pumping and the transition from seismic slip to ductile creep are both the result of the interplay of fluid pressure and deformation mechanisms (Irwin and Barnes, 1975; Sibson et al., 1975; Sibson, 1990. In recent years, deep drilling of the continental crust has revealed the presence of fluids at depths in excess of 12 km.

Oxygen and hydrogen isotope studies of fluid processes in the crust have been applied largely in the context of tracing fluid sources and geothermometry. Stable isotopes have also been employed to test for equilibrium versus disequilibrium, for open versus closed system conditions, and for estimating fluid/rocl ratios. Combined radiogenic and stable isotope determinations have been successful in resolving the conditions under which radiometric systems persist or are disturbed (Frey et al., 1976; Kerrich, 1991), and this approach has also been adopted for testing resetting of paleomagnetic signatures (Ellwood and Wenner, 1981). The geochemical and stable isotope systematics of fluids has been dealt with in a number of specialized and general papers. Specifically, the stable isotope systematics of hydrothermal fluids in ore deposits have been reviewed by Taylor (1974; 1979) for oxygen and hydrogen; by Ohmoto and Rye (1979) for carbon and sulfur; and by Ohmoto (1986) for H, C, O and S-isotopes.

Numerous papers have addressed the subject of fluid flow through the crust with particular emphasis on oceanic ridge systems and continental magmatic centres. However, few papers have dealt with advection of fluids in other tectonically active areas. This overview summarizes the results of geochemical, stable and radiogenic isotope studies of fluid processes in the crust, with emphasis on extensional, collisional and transpressional geodynamic regimes.

*Classification of fluids*

Fluids interacting with the crust can have a variety of compositions. Here, we adopt the following definitions modified from Kharaka et al., (1985) for fluids based on their chemical compositions:

fresh water has < 1000 mg/l solutes
brackish water has 1000-10,000 mg/l solutes
saline water has 10,000-100,000 mg/l solutes
brine has > 100,000 mg/l solutes

Reference is often made in the literature to various types of waters, and some of these have genetic connotations whereas others do not. Formation water refers simply to the fluid resident in rocks, and has no significance to origin or age. Connate water is the fluid deposited with the sediments or rocks and can be modified via reactions with the reservoir rocks. Meteoric water originates from rain or snow and can be modified via interaction with rocks at elevated temperatures to become a hydrothermal fluid. Metamorphic water is that associated with metamorphic dehydration reactions. Magmatic water is that hydrous fluid released from a crystallizing magma. Other than formations waters, these "categories" of water are applicable to describing the origin of only some fluids because the extensive interactions between most fluids and rocks obscures their origin. The observed ranges in the isotopic composition of fluids that may affect the crust are shown in Figure 1, where it is clear that the different types of fluids referred to above also have distinct H and O isotope compositions. Fluids may shift in isotopic composition by reaction with rocks. For example, meteoric water or seawater may evolve isotopically into the "metamorphic" field. When seawater evaporates, the $\delta$ and $\delta^{18}O$ values of residual brines first increase as evaporation proceeds, and then decrease when the solute concentration becomes high enough to reverse the fractionation factor between water liquid and water vapor (Sofer and Gat, 1975). Fluids from Precambrian shields, such as Sudbury and Yellowknife of the Canadian Shield, have peculiar isotopic compositions that plot the left of the meteoric water line.

Inasmuch as water is comprised only of hydrogen and oxygen, $\delta D$ and $\delta^{18}O$ values should be valuable tools for tracing the origin and evolution of a fluid. Chemical compositions may also be used to trace the evolution of a fluid, but rarely does any constituent other than hydrogen (and occasionally oxygen) isotopes behave most frequently as conservative elements that can be used to trace the origin of a fluid.

## ACCRETIONARY PRISMS AND SUBDUCTION ZONES

The transport rate of water from the hydrosphere into subduction zones is estimated at $9.7 \pm 3.4 \times 10^{14}$ g $a^{-1}$ (Fyfe, 1978; Ito et al., 1983). This estimate takes into account $H_2O$ as a structural component of hydroxy/silicates in altered oceanic crust and its sedimentary cover, but not the pore water content of such rocks which may be large. According to Fyfe (1978) and Ito et al., (1983) the reflux of subducted $H_2O$ to the surface cannot be accounted for by a combination of arc, spreading centre, and hot spot magmatism, yielding an apparent disparity in the global mass balance of $H_2O$ equivalent to a net flux of $1.2\text{-}11 \times 10^{14}$ g $a^{-1}$ into the mantle. The net flux of carbon into the mantle is of similar magnitude.

At convergent plate margins, oceanic sediments are deformed as they are scraped from the subducting oceanic lithosphere and form an accretionary prism, schematic example of which is shown in Figure 2. Pore fluids are expelled from sediments during deformation even deep in the complex where dehydration of minerals and hydrocarbon generation can occur. The porosity of the sediments along trench axes typically is 40-80% but can be reduced to 20% during accretion (Bray and Kavig, 1985) during which fluids can escape via a multitude of pathways. Expulsion may occur by (1) permeation vertically through the subduction complex, although low permeabilities of many units and the chemical composition of seawater above the complex suggest this effect is minimal, (2) flow updip along permeable strata (Kulm et al., 1986; Ritger et al., 1987), or (3) migration along permeable faults associated with deformation of sediments, especially focused at the decollement, which is the convergent plate boundary (Davis et al., 1983; Vrolijk, 1987; Moore et al., 1987). From those considerations of fluid mass balance it is clear that large volumes of thermal waters must advect along continental margin faults above shallow subduction zones and along the decollement and related low angle thrust faults in accretionary prisms.

*Accretionary prisms*

Many factors appear to influence the sources and pathways of fluids in convergent margins. The overall structures of accretionary wedges appear to be a complex function of the dip of the subducting slab (Davis and Hyndman, in press), type of sediments in the wedge (Moore, 1975; Karig and Sharman, 1975; Vrolijk, in press), and rate of convergence (Elderfield et al., in press).

Along the North Cascadia subduction zone, Davis and Hyndman, (in press) used seismic reflection profiles to suggest that the dip of 11°C for the Juan de Fuca plate off Vancouver Island produces the Tofino sedimentary basin rather than upward growth of an accretionary prism in accordance with the model of Davis et al., (1983), whereas to the southeast where the subducting slab dips more gently, the accretionary prism grows upward to form the Olympic Mountains. Near hydrostatic pressures and steeply dipping faults occur in turbidites in the section while deeper and less permeable hemipelagic sediments have pore fluid pressures that approach lithostatic. Landward-dipping thrust faults penetrate the entire sedimentary section and little if any of the sediments appear to have been subducted since the Eocene, as they are contained in the accretionary prism.

Evidence from accretionary prisms along continental margins indicates that pelagic sediments appear to be subducted and terrigenous trench deposits are accreted at the toes of prisms (e.g. Moore, 1975; Karig and Sharman, 1975). This may result because of the differential changes in the mechanical properties of sediments during subduction such that detachments of trench deposits from pelagic deposits occur at the less competent hemipelagic sediments. In other accretionary prisms, such as the Barbados accretionary complex, the decollment occurs at the base of the smectite-rich section and may be controlled by the weakest stratigraphic horizon. Vrolijk, (in press) has suggested that decollement zones form in smectite-rich horizons in which pore fluid pressures are less than lithostatic pressures and that the smectite to illite transformation occurs at the depth of the first earthquakes in subduction zones.

Localized fluid flow presently occurs in accretionary prisms as evidenced by seeps, mud volcanoes and diapirs, unique biological communities, and chemically and thermally anomalous fluids issuing from vents or within the sediments of the prism (Langseth and Moore, in press). Subduction-related venting of fluids has been observed in the Oregon-Washington Accretionary Complex (Moore et al., 1984; Suess and Massoth, 1984; Kulm et al., 1986), the Barbados Trench (Moore et al., 1987; Gieskes et al., 1990), the Naikai Trough and Japan French (Boulegue, et al., 1987), and the Peru Margin (Kulm et al., 1986; Suess et al., 1988). Fluid flow in deeper portions of the subduction zone have been documented in the exposed Kodiak Accretionary Complex of Alaska by Vrolijk, (1987) and Vrolijk et al., (1988) who examined coexisting methane-rich and water-rich fluid inclusions in syntectonic quartz veins.

The Northern Barbados Accretionary Complex represents a slow-convergent tectonic environment (~2 cm/yr) where a broad wedge of sedimentary material has accumulated from subduction of the North American plate (Moore et al., 1987; Mascle et al., 1988). Core recovered from ODP Leg 110 near Barbados reveals a complex structure with the decollment zone existing at the most smectite-rich part of the stratigraphy (Wilkens et al., 1989; Tribble, in press). Anomalous heat flow values of 80-300% higher than normal oceanic crust of the same age occur at 15°30'N on the edge of this convergent margin and decrease downhole, suggesting that the source of heat is in the sediments (Fisher and Hounslow, in press). The heat flow results are consistent with episodic flow of an average of $10^{-7}$ m/sec of warm fluids along thrust faults, sandy intervals, and the major decollment zone where permeabilities are enhanced.

Comparison of the trends in the chemical and isotopic compositions of normal oceanic pore fluids with the composition of fluids in the sediments of the accretionary prism at Barbados reveals that in the highly faulted areas and near the decollement zone, the fluids are depleted in chloride and enriched in $^{18}O$, $SO_4$, $CH_4$, and perhaps deuterium (Figure 3). Vrolijk et al., (in press) have interpreted the trends at Barbados to reflect alteration of volcanic ash to smectite within the sediments and underlying basalt coupled with diffusive exchange with seawater, as normally occurs in oceanic sediments, but complicated by fluid migration from deep within the prism along faults.

Dewatering of smectite interlayers and the conversion of smectite to illite deep beneath the accretionary prism may be the sources of the $^{18}O$-rich and chloride-improverished fluids. Thermal cracking of organic matter at depth may be the source of the methane in the Barbados prism, which is unlike other convergent margins such as the Peru Trench where hydrates occur and are probably sources for methane, $^{18}O$ and deuterium (Shipley et al., 1979; Kvenvolden and McManamine, 1980). Low chloride and high methane and $^{18}O$ contents of fluids released in accretionary prisms appear to be common.

High methane and low chloride contents of syntectonic fluids generated in subduction zones are also recorded in the fluid inclusions in syntectonic quartz veins in the presently exposed Kodiak Accretionary Complex, Alaska (Vrolijk, 1987). Water-rich inclusions in these veins yield trapping temperatures of 215-290°C and methane-rich inclusions have lower densities towards the centers of the veins. Changes in the densities of methane at constant

trapping temperatures imply the fluids were trapped at pressures equal to 10-14 km depth which later decreased in the fractures as the fluid pressures became less than lithostatic.

Although studies of accretionary prisms are limited, it is becoming clear that there is flow of warm fluids having distinct chemical and isotopic compositions back into oceans as a result of tectonic activity at convergent margins, especially along decollment zones. The flux of this fluid is difficult to estimate at present because it is also evident that the migration and mixing histories are complex processes.

*Continental margin faults*

Vancouver Island is located in a region of active seismicity, above the eastward subducting Juan de Fuca plate. The island is part of an allochthonous terrane accreted onto the N. American continent. It is transected by numerous large scale faults many of which have recorded significant displacements as recently as the Eocene (Clowes and Brandon, 1985).

Several of the major faults, such as the San Juan and Taylor River faults, which transect greenschist facies rocks, are sporadically mineralized with Mn-oxides, or hydrocarbon-carbonate fracture-infilling. $\delta^{13}C$ values of the carbonate range from -14 to 4 per mil, and $\delta^{18}O$ values from 3 to 25 per mil. Fyfe et al., (1987) interpreted the mineralization and isotopic data in terms of distillation of volatiles and solutes from hydrocarbon-bearing sediments on the Juan de Fuca plate during subduction. The $^{18}O$ depleted carbonates may reflect isotopic overprinting by recent meteoric waters, or mixing of such waters with subduction fluids advecting up the faults.

## BRINES IN PRECAMBRIAN SHIELDS

Brines in crystalline rocks of Precambrian Shields have been a nuisance to miners and engineers and an unsolved mystery to geologists and geochemists. Intermittent discharge of brines have been reported from several mining districts, especially in Canada, as well as in deep drill holes in waste disposal programs. In addition to brines, fractures filled with secondary minerals deposited from migrating fluids are common even in the deepest drill holes. Much of the data on brines and gases in crystalline rocks are summarized in Fritz and Frape, (1987).

The vast majority of studies on brines in crystalline rocks involve those in the Precambrian Shield of Canada. Thirty-two occurrences of brine or saline water have been documented, although the vast majority of data are from mines at Yellowknife, Thompson, Sudbury, and Matagami. Brine and saline fluids collected from these areas have similar chemical characteristics and can be classified as Ca-Na-Cl solutions. The most pronounced chemical characteristic of the brines is that they lie to the left of the meteoric water line on a $\delta D$ versus $\delta^{18}O$ plot and are isotopically distinct from seawater or basinal brines (Figure 1, 4). When all of the fluids from a locality are considered, the total range in hydrogen and oxygen isotopic compositions can be modelled as mixing of modern meteoric waters in each area (as represented by fresh and brackish waters) with $^{18}O$- and deuterium-rich brines having $\delta D$ and $\delta^{18}O$ values that plot left of the meteoric water line (Frape and Fritz, 1987). Frape and Fritz, (1982; 1987) extrapolated the $\delta D$ and $\delta^{18}O$ values of fluids from the

Canadian Shield and concluded that the trends from all distinct areas intersected near the same $\delta D$ and $\delta^{18}O$ values, so that the brine end member throughout the Shield may have had the same isotopic composition. The $\delta D$ and $\delta^{18}O$ value of the brine end member obtained in this way is -20 to 0 and -10 to -7 per mil, respectively, similar to the values obtained by Pearson, (1987) who extrapolated the isotopic compositions of the fluids to halite saturation. If there is a brine end member that has the same D/H and $^{18}O/^{16}O$ ratio throughout the Canadian Shield, its consistency would require the same origin or a similar process to alter any original variations.

The chemical compositions of fluids from the Canadian Shield vary with the depth they are collected and, hence permeability, such that samples from the deepest levels in the Shield have the highest total dissolved solids. Frape and Fritz, (1987) found that chloride and $Ca^{+2}$ also increase with depth, especially deeper than 1000m, and that all the brines have <20 tritium units. In contrast, $\delta D$ values do not vary regularly with depth suggesting that the processes that control the isotopic compositions probably do not control the chemical compositions and vice versa. Except for samples collected from depths less than 200m having low TDS contents, $Ca^{+2}$, $Sr^{+2}$, and bromide contents correlate with chloride contents but $Mg^{+2}$, $K^+$, B, fluoride, and iodide contents do not. None of the relations between any of the constituents and chloride contents are compatible with the compositions of seawater, evaporated seawater, or mixtures of seawater and freshwater (Frape and Fritz, 1987).

There are two major theories about the origin of the brines in Precambrian Shields. They are (1) the brines represent modified Paleozoic seawater or basinal brines (Kelly et al., 1986; Guha and Kanwar, 1987), and (2) the brines are produced by leaching of saline fluid inclusions in crystalline rocks or by intense rock-water interactions (Nordstrom and Olsson, 1987; Frape and Fritz, 1987; Gascoyne et al., 1987; Kamineni et al., 1987). Equilibrium phase diagrams in the system $K_2O$-$Na_2O$-CaO-$SiO_2$ indicate that the chemical compositions of the fluids in the Precambrian Shield of Canada are controlled by minerals similar to those in rocks which host the brines (Figure 5). For example, the relation between the activities of $Na^+/H^+$ and $K^+/H^+$ suggests that the chemical compositions of the Shield fluids are distinct from seawater and basinal brines and are controlled by kaolinite and muscovite (and feldspar in some samples) at ~25°C. Activities of $Na^+/H^+$ and $H_4SiO_4$ are compatible with equilibrium with kaolinite, smectite, and albite at low temperatures.

Kaolinite and other clay minerals probably control activity relations between $K^+/H^+$ and $H_2SiO_4$, whereas kaolinite, smectite, and laumontite may control $Ca^{+2}$, especially at Sudbury. The temperature of interaction must be low otherwise the fluids would be enriched in $^{18}O$ as are the vast majority of hydrothermal fluids. Thus, concentrations of most of the major cations are in equilibrium with kaolinite, smectite, and muscovite (and laumontite in some samples), all of which are present in the granites and mafic meta-igneous rocks which host the fluids at Yellowknife and Sudbury. Further, the range in activities cannot be attributed to dilution of an end member brine with freshwater, as is suggested by the hydrogen and oxygen isotope systematics, but probably results from interactions with rocks having different lithologies. Pearson, (1987) also suggested that the Na/K ratios in the brines from the Canadian Shield are controlled by equilibrium reactions with alkali feldspars and muscovite

and Ca/Na ratios by plagioclase at low temperatures. Frape et al., (1984) also concluded that the salinities of the brines result from intense water-rock interaction with local hosts but at unspecified temperatures. McNutt et al., (1990) reported substantial ranges of 0.704 to 0.753 for the $^{87}Sr/^{86}Sr$ ratios in brines and fracture-filling minerals from the Canadian Shield, which they attribute to interaction of the fluids with feldspars in the local rocks.

Although the chemical compositions of the fluids in the Canadian Shield probably result from extensive water-rock interactions at low temperatures, the reasons for the hydrogen and oxygen isotopic compositions of the brines are less obvious. The possible processes that could result in isotopic compositions to the left of the meteoric water line include: (1) precipitation of substantial quantities of hydrous minerals having high $\delta^{18}O$ values and low $\delta D$ values at low temperature from limited amounts of seawater or meteoric waters; (2) preferential exchange of hydrogen isotopes in meteoric waters having low D/H ratios with hydrous minerals in the rocks and fractures having much higher D/H ratios; and (3) loss of hydrogen from the fluids as a result of the reduction of water to deuterium-depleted methane or reduced hydrogenous gases which enrich the remaining water in deuterium.

Processes involving both hydrogen and oxygen isotope exchange with minerals or formation of the fluids from fluid inclusions, such as is suggested by the chemical compositions of the brines, should result in trends similar to those observed in hydrothermal fluids or basinal brines and therefore cannot explain the brines that plot to the left of the meteoric water line. Precipitation of hydrous minerals, which would be enriched in $^{18}O$ by ca. 30 per mil and in deuterium by 50 per mil relative to the brine, could in theory move a fluid to the left of the meteoric water line. However, the fluid/mineral ratio would have to be prohibitively low, i.e. >50% of the original mass of the fluid must go into hydrous minerals. This value is a minimum and likely to be higher because exchange of isotopes between the fluid and minerals, which would increase the $\delta^{18}O$ value of the fluid, is necessitated by the concentration of cations required to precipitate the hydrous minerals. Although preferential exchange of hydrogen isotopes between fluids and hydrous minerals at low temperatures is known to occur (Wilson et al., 1987; Kyser and O'Hanley, 1989; Bird and Chivas, 1989), the amount of hydrogen in the fluid is so large relative to that in hydrous minerals that unreasonably low fluid/rock ratios (i.e. <.001) are required to affect only the $\delta D$ value of the fluid.

The most plausible mechanism to explain the peculiar isotopic compositions of the brines in the Canadian Shield would be the preferential loss of H relative to deuterium via a very D-depleted gas. Fritz et al., (1987) reported that methane associated with brines from the Canadian Shield have $\delta D$ values between -500 and -100 per mil. Although they favor an origin by abiotic processes such as hydrothermal alteration or low-temperature kinetic processes, this methane is associated with the brines and serves as a sink for hydrogen in the system. Such a process would explain why the $\delta D$ values and chemical compositions of the brines may be decoupled in many samples.

*Fracture-filling minerals*

Multiple generations of fracture-fillings have been encountered in a variety of plutonic rocks during deep drilling in the Canadian Shield. These form a systematic sequence of

diminishing temperature and relative age, from pegmatites and aplites, to epidote-chlorite, adularia, prehnite, laumontite-calcite, and clay (kaolinite, palygorskite, halloysite) fracture-fillings (Kamineni and Stone, 1983 and references therein). Upper plateau $^{40}Ar/^{39}Ar$ ages of fracture generations that are amenable to dating range from Archean (2,650 Ma) to 960 Ma.

Combined stable isotope and fluid inclusion data on the fracture filling minerals has revealed a complex and protracted hydrologic history of the Shield, spanning 2,700 Ma. A generation of epidote-chlorite veins formed at 2,650 +/- 15 Ma, over the temperature interval 160-280°C, from saline NaCl-$CaCl_2$ brine having $\delta^{18}O$ values of 0.2 to 3.0 and $\delta D$ values of -20 to 0 per mil. These thermal brines have chemical and isotopic compositions compatible with those of evolved Archean marine water that were likely convected during emplacement of late Archean granites.

Successive generations of fracture arrays formed at 1100 Ma, 960 Ma, and in the Lower Paleozoic and are tentatively linked to Keweenawan and Grenvillian activity, and the onset of sedimentation of the Canadian Shield in the Cambrian, respectively. The sequential fracture-fillings record progressively diminishing temperatures and the involvement of more deuterium- and $^{18}O$-depleted fluids. For example, adularia-chlorite veins, formed at 960 Ma, from fluids having $\delta^{18}O$ values of -1 to -7 per mil and $\delta D$ values of 25 +/- 5 per mil. Fluid inclusions in all generations of veins record variably saline (7 to 30 wt.% NaCl equivalent) $CaCl_2$-NaCl brines. Narrow clay-lined fractures are the site for seepage of modern groundwaters. The clay isotopic signatures fall on the kaolinite line confirming their formation (or re-equilibration) in the presence of recent near-surface waters (Figure 4).

Collectively, these results from fracture arrays in the Superior Province imply that fluids infiltrated Precambrian rocks of the Shield over a protracted time span, and underwent chemical and isotopic evolution by promoting the formation of secondary minerals. Fluids were intermittently expelled along reactivated faults and fractures probably by seismic pumping, during magmatic and tectonic events that disturbed the Shield (Sibson et al., 1975; Sibson, 1981; Kerrich and Kamineni, 1987). For a normal continental geotherm of 25°C/km, and fracture-filling temperatures of 250°C, the fluids must have penetrated to depths of at least 10 km.

## FLUIDS ASSOCIATED WITH COMPRESSIONAL TECTONIC REGIMES

### *Alpine belt*

Petrologic studies in the European Alpine belt have revealed pronounced differences between the basement and cover tectonic units, in terms of metamorphic temperatures, effects of overprinting, scale of isotope exchange, and fluid regimes. In general, the basement units (Penninic nappes, or thrust sheets) are pre-Mesozoic in age and polymetamorphic, whereas the Mesozoic cover has experienced only Alpine tectonism and metamorphism. Stable isotope studies have been conducted on the Swiss Central Alps (Hoernes and Friedrichsen, 1980), the Monte Rosa area (Frey et al., 1976), the Otztal-Stubai Alps, Austria (Hoernes and Friedrichsen, 1978), the Helvetic Alps (Burkhard and Kerrich, 1988), and the Sesia Zone, Italian Alps (Desmons and O'Neil, 1978).

A great variety of fracture filling types are present in both basement and cover tectonic units of the Alpine orogenic belt. These include tensile and shear fractures, compound fibrous veins (Durney and Ramsay, 1974), and pressure shadows. Some of the veins are syntectonic with respect to the main Alpine deformation, and may contribute a significant component of the bulk strain, whereas other fracture arrays appear to be post tectonic. Mineral parageneses indicate that fracturing operated over a wide range of pressure-temperature conditions, from subgreenschist facies to within the field of sillimanite stability. Light stable isotope studies of Alpine fracture systems have been undertaken in order to: (1) map out ambient temperatures of brittle deformation using isotope geothermometry, (2) test for open versus closed system behavior during fracture filling, (3) evaluate the scale of isotopic exchange between different lithological units, and (4) identify the origin of fluids present during fracture propagation-infilling.

*Basement*

The Penninic basement nappes are characterized by relatively uniform $\delta^{18}O$ values of minerals in gneisses of contrasting protolith, such as amphibolites, metapelites, and ortho- and paragneisses, where primary $\delta^{18}O$ differences in excess of 10 per mil would be anticipated. Hoernes and Friedrichsen, (1978; 1980) attributed this uniformity to large scale isotopic homogenization during a pre-Alpine metamorphic event. Concordant isotopic equilibrium is preserved between quartz, biotite, muscovite and ilmenite in metapelites, and between quartz, garnet and rutile in eclogites, reflecting this pre-Alpine metamorphism.

In the Monte Rosa area, concordant isotopic temperatures of 520 to 560°C are preserved among quartz, muscovite, and biotite, from the Permian metamorphism and Permian Rb-Sr isochron ages are also retained. These isotopic features are commensurate with the persistence of pre-Alpine metamorphic assemblages, and the absence of retrograde cation exchange reactions. Frey et al., (1976) concluded that the preservation of Permian isotopic and mineral equilibria is due to closed system behavior during Alpine orogenic events.

Desmons and O'Neil (1978) also identified closed system behavior in Austro-Alpine units from the Sesia Zone, during Alpine deformation. The Austro-Alpine units are comprised of a diverse lithological ensemble including eclogites, glaucophone schists, and marbles having a polymetamorphic history. In the majority of rocks, the $\delta^{18}O$ values of specific minerals are uniform, due to efficient isotope homogenization, although eclogites tectonically emplaced in marbles are $^{18}O$ enriched owing to exchange with the higher $^{18}O$ carbonates. This isotopic exchange with $^{18}O$ rich carbonates occurred during a Hercynian or Eoalpine metamorphism, in which the high pressure phases developed. Sesia Zone eclogites are isotopically comparable to those from Monte Rosa (Frey et al., 1976) and California type IV eclogites, but are $^{18}O$ enriched relative to type C eclogites from Zermatt, Venezuela, New Caledonia and Tiburan (Taylor and Coleman, 1968; Vogel and Garlick, 1970; Black, 1974). Isotopically concordant temperatures of 540 +/- 20°C are preserved among quartz, rutile, and phengite. Desmons and O'Neil (1978) concluded that fluids were internally buffered with respect to individual lithotypes (in contrast to the widespread isotopic exchange observed in the Austrian basement units), but that pervasive hydrogen isotope exchange occurred throughout the sequence in the pre-Alpine metamorphism (Figure 6).

Major ductile shear zones transect several of the basement granitoid 'massifs' in the Alpine belt. At Mieville, in the Aguilles Range massif, a large ductile shear zone in granodiorite apparently formed under isochemical conditions. Whole rock $\delta^{18}O$ values, $Fe^{2+}/\Sigma Fe$ ratios, and major and trace element abundances, are remarkably uniform from undeformed protolith to the most intensely strained tectonite (Kerrich et al., 1987). This is consistent with the generally closed system behavior in basement rocks during Alpine deformation.

*Alpine fissures*

The so-called Alpine fissures are a distinctive generation of tensile fracture fillings occupied by quartz, carbonates, and feldspars, with subordinate biotite and chlorite. Apatite, epidote, diopside and beryl have also been reported (Luckscheiter and Morteani, 1980). This fracture set post-dates the main Alpine tectonism, and was generated during Alpine metamorphism in an environment of uplift and doming.

In Penninic gneisses of the Tauern Window (Austria) there is a compliance in the $\delta^{18}O$ values of quartz and biotite between host rock and fissure. Fissure quartz in amphibolites having $\delta^{18}O$ values near 5 per mil has $\delta^{18}O$ values of (~6 per mil), which is lower than counterparts ($\delta^{18}O$ quartz ~ 10 per mil) in granite gneiss. Friedrichsen and Morteani, (1979) concluded that individual rock-fissure systems are internally buffered with respect to oxygen isotopes. The total range of temperatures, calculated from quartz-biotite is 640 to 450°C, with a maximum in the central Tauern Window area, diminishing systematically to the north and south, a distribution representing deformation of predoming isothermal surfaces.

Isotopic buffering of fluids at the scale of individual lithological units is also reflected in the composition of fluid inclusions within Alpine fissures. Luckscheiter and Morteani, (1980) reported that fluid inclusions in eclogites contain a high density $CO_2$ phase trapped at ~7.5 kb, inclusions in schists are characterized by high $CO_2/H_2O$ ratios but low salt contents, in the Zentralgneis fluid inclusions are low $CO_2/H_2O$ but with enhanced salt contents, and calc-mica schists have inclusions with low salt contents and no detectable $CO_2$.

*Alpine Overprinting*

Localized effects of Alpine overprinting on basement gneisses have been identified. Maximum Alpine metamorphic temperatures in the Monte Rosa area were ~460°C (Frey et al., 1976). Isotopic reversals occur sporadically, where $\delta^{18}O$ values of albite > K-feldspar, and $\delta^{18}O$ muscovite > phengite. The former has been accounted for in terms of albite precipitation during Alpine times from a fluid with which K-feldspar did not exchange. Some biotites retain isotopic temperatures of Permian metamorphism, but have K-Ar systems reset to Alpine ages. Pre-Alpine mineral parageneses, together with radiometric and stable isotope equilibria have also been preserved in regions of the Aar and Gothard massifs which have escaped Alpine deformation (Hoernes and Friedrichsen, 1980).

In the Swiss Central Alps, the maximum temperature of Alpine metamorphism exceeded 600°C, but not all oxygen isotope fractionations in the Pennine basement units have not adjusted to this maximum. Quartz-garnet and quartz-ilmenite isotope temperatures are systematically larger than those derived from quartz-biotite. Hoernes and Friedrichsen,

(1980) argued that the quartz-biotite temperatures were obtained during prograde metamorphism, and prior to the maximum temperature, owing to the cessation of cation exchange reactions and tectonic recrystallization involving biotite. This effect is also seen in the Otztal-Stubai Alps, where quartz-ilmenite fractionations correspond to temperatures >600°C, whereas quartz-biotite and quartz-muscovite yield temperatures of ~500°C. Alternatively, biotite may have undergone preferential exchange during post peak metamorphic cooling, relative to garnet and ilmenite, such that quartz-biotite fractionations reflect retrograde rather than prograde conditions. This effect has been documented in metamorphic rocks of the Limpopo belt by Huebner et al., (1986).

*Cover rocks in the Alps*

Mesozoic metasedimentary rocks of the tectonic cover are isotopically more variable, and generally $^{18}O$ enriched, relative to the basement gneisses. $\delta^{18}O$ values of quartz span 7 to 21 per mil, and correlate with lithotype, such that some isotopic memory of the sedimentary precursor is retained. Baertschi (1957) reported that Mesozoic carbonates preserve their original sedimentary isotopic composition where juxtaposed with basement gneisses and show no correlation of $\delta^{18}O$ with metamorphic grade. Quartz, muscovite and biotite generally preserve triple oxygen isotopic concordancy that reflects Alpine metamorphism. Muscovite is sporadically reset to lower $\delta^{18}O$ values, and quartz-carbonate fractionations often indicate disequilibria.

The relative sensitivities of O and C-isotope signatures of carbonates to resetting, is illustrated in a study of Jurassic oceanic crust subject to progressive intensities of Alpine tectonism, in the southern Pennine thrusts of Eastern Switzerland (Weissert and Bernoulli, 1984). Carbonates in units that have undergone the highest intensities of deformation and metamorphism, have $\delta^{18}O$ values of calcites that are depleted 5 to 6 per mil relative to unaffected precursors, whereas $\delta^{13}C$ values remain consistent at -1 to +2.5 per mil (PDB).

*Hydrogen isotope systematics*

Biotite and muscovite in the pre-Alpine basement gneisses are characterized by a restricted range of $\delta D$ values, independent of rock type. Biotite $\delta D$ values cluster at -35 to -60 per mil in the Otztal-Stubai Alps, at -40 to -80 per mil in the Monte Rosa area, and -60 to -65 per mil in the Swiss Central Alps (although the total range here is -50 to -140 per mil). Pre-Alpine phengites in the Sesia Zone are confined to between -30 and -45 per mil. Hoernes and Friedrichsen, (1978; 1980) attributed the narrow range of $\delta D$ values to pervasive exchange with magmatic fluids in a pre-Alpine metamorphic event. However, $\delta D$ values of biotite in equilibrium with 'magmatic fluids' would be significantly more depleted (-80 to -110 per mil) than the observed values. An alternative explanation for $\delta D$ values as enriched as -35 per mil was proposed by Frey et al., (1976), who suggested that deuterium enriched connate fluids originally affected the rocks. The presence of such fluids would also account for the widespread oxygen-isotope homogenization observed in the Austrian basement gneisses (see above). Any mechanism for the D-enriched biotites must account for their presence in basement gneisses throughout the entire Western and Eastern Alps.

In Mesozoic metasedimentary cover rocks, $\delta D$ values of biotites are inherently more variable, and extend to lower absolute values, than in the basement gneisses. For example,

the δ D values of biotite = -40 to -95 per mil in the Otztal-Stubai Alps, -60 to -130 per mil in the Swiss Central Alps, and -134 to -157 per mil at Monte Rosa. Two processes have been proposed to explain the populations of most deuterium depleted biotites (δ D = -80 to -157 per mil): (1) Infiltration of an external deuterium depleted aqueous reservoir during Alpine metamorphism (Frey et al., 1976; Hoernes and Friedrichsen, 1978; 1980), and (2) Loss of water accompanying metamorphic recrystallization (Hoernes and Friedrichsen, 1978). In this scheme, residual hydroxyl phases are postulated to have become depleted in deuterium during dehydration in an open system, owing to the magnitude of negative mineral-water DH fractionation.

There are difficulties in accounting for the most depleted biotites (~-140 per mil) by the second mechanism, given that mineral-water fractionations are less than 60 per mil at all temperatures in excess of 500°C (Suzuoki and Epstein, 1976). The contrast between uniform biotite δ D values throughout the Penninic basement gneisses (δ D = -57 $\pm$ 3 per mil), and larger dispersions in the Mesozoic cover (δ D spans -60 to -90 per mil), rules out extensive fluid exchange between basement and cover during Alpine tectonism. δ D and δ $^{18}$O values of biotites and muscovites are not correlated in the population of low-δ D micas, but record only a δ D shift. Plausibly the micas have preferentially undergone H-isotope exchange with meteoric water at relatively low temperatures, whilst at the same time remaining far from isotopic equilibrium with oxygen in the water. This interpretation is consistent with evidence from calcite-filled fractures for infiltration of low δ $^{18}$O fluids during buoyant rise of the Alpine mountain belt (Figure 7), and experimental studies showing that the rate of H diffusion in micas is faster than for oxygen, although some uncertainties are associated with interpretation of the experimental results (Graham, 1981). Because isotopic exchange does not normally occur in the absence of cation or anion reactions (Kyser, 1987; O'Neil, 1987), it would be interesting to compare the OH/F partitioning between the populations of 'normal' and low δ D biotites. Calculated isotopic compositions of water in equilibrium with the basement and cover units are illustrated in Figure 6.

*Pyrenees*

A similar situation to the H and O-isotopic uniformity in pre-Alpine basement has been described by Wickham and Taylor (1985) from the Trois Seigneurs Massif in the French Pyrenees. This region underwent high temperature and low pressure metamorphism in the Hercynian orogeny (Late Carboniferous). All metamorphic rocks above the 'andalusite in' isograde form a relatively homogeneous population of whole rock and quartz δ $^{18}$O values, including metacarbonates, metapelites, and granitic rocks, whereas lower grade counterparts to the metapelites are $^{18}$O enriched (δ $^{18}$O ~15 per mil). δ D values of muscovites are tightly clustered and high (δ D = -25 to -35 per mil), as is the case for micas in the pre-Alpine basement.

Wickham and Taylor, (1985) accounted for the combined isotopic shift and homogenization by pervasive incursion of aqueous fluids with an externally buffered composition, probably seawater. Simple convective circulation of fluids flattens, rather than steepens, thermal gradients, and requires high permeability from deep fractures. These conditions contrast with the steep thermal gradient and ductile deformation in the high grade to antectic environment. In order to circumvent these difficulties, Wickham and

Taylor (1985) suggested that convective circulation of marine water took place just prior to the metamorphic maximum. They rule out magmatic or metamorphic water as the external reservoir because $\delta\,^{18}O_{H2O}$ values in the range of 4 to 8 per mil would require water/rock ratios in excess of 0.3 by mass, which were believed to be unreasonably large. Meteoric water was ruled out by the high $\delta$ D values of the micas, and connate fluids were regarded as an implausible fluid source if the observed isotopic homogenization existed throughout the Pyrenean basement.

Significant isotopic gradients are known to exist in some sedimentary sequences where fluids having the highest $\delta\,^{18}O$ and $\delta$ D values are usually the deepest. Accordingly it may be unnecessary to invoke marine water incursion in the Trois Seigneurs massif in order to account for the observed isotopic uniformity and high $\delta$ D values. For example, Eslinger and Savin (1973) have shown that in the Belt Supergroup metasediments, the $\delta\,^{18}O$ values of whole rock decrease systematically with respect to depth. Over a vertical distance of 2 km, $\delta\,^{18}O$ values of pelitic lithologies diminish from 18-19 per mil down to 12-14.5 per mil, and over the same interval carbonates are shifted by -7 per mil. Similar trends have been monitored in the Gulf Coast sedimentary basin, where bulk silicates shift from 19.6 to 18.0 per mil and carbonates from 26.3 to 23.1 per mil over a stratigraphic depth of 4 km: over the same interval fine-grained clays show depletion of 4.5 per mil (Yeh and Savin, 1977). Trends in $^{18}O$ with respect to increasing stratigraphic depth are mirrored by an enrichment in deuterium of clay minerals, from -73 to -33 per mil (Yeh, 1980). At the deepest levels, clays are characterized by $\delta$ D values that match those of micas in the Hercynian basement of the Alpine and Pyrenean belts.

Patterns of $^{18}O$ depletion, isotopic homogenization, and deuterium rich micas are widespread in Hercynian basement of the Alpine and possibly Pyrenean belts. Whether or not marine water was involved in all of these areas, or alternatively connate fluids, or tectonically driven brines (Fyfe and Kerrich, 1985; Oliver, 1986), are questions that remain to be resolved.

*Thrust faults in the Western U.S.A.*

Infiltration of meteoric waters along major thrust faults has been identified at a number of localities in the Western States. Lee et al., (1984) have shown that the $\delta$ D and $\delta\,^{18}O$ values and apparent K-Ar ages of Mesozoic and Tertiary plutons in Nevada have been disturbed in proximity to the Snake Range decollement. In Oligocene plutons, $\delta$ D and $\delta\,^{18}O$ values, and K-Ar ages are as low as -2.5 per mil, -155 per mil, and 18 Ma, respectively, in zones of deformation where enhanced hydraulic conductivity has facilitated the incursion of isotopically light thermal waters. For undeformed Jurassic plutons the $\delta\,^{18}O$ values are undisturbed, but both the $\delta$ D and K-Ar ages have been perturbed over distances of tens of metres below the decollement. This disparity in the sensitivity to resetting of O, H and K-Ar isotopic systems, parallels the trends observed in the Alpine belt, as previously discussed.

## FLUIDS ASSOCIATED WITH EXTENSIONAL TECTONIC SETTINGS

Metamorphic Core Complexes constitute a distinct structural-lithological entity in the tectonic architecture of the North American Cordillera. They are geologically characterized by a 'core' of intensely deformed metamorphic and (or) plutonic rocks, grading upward

through gently dipping mylonites to a low-angle detachment fault. This infrastructure, or lower plate, is overlain by a suprastructure, or upper plate, of attenuated, listrically faulted, allochthonous, and generally unmetamorphosed Tertiary rocks. Core complex development postdated Mesozoic (Laramide) thrusting, preceded basin-range faulting, and was broadly contemporaneous with the eruption of Tertiary ignimbrites. Twenty-five such complexes have been identified to date, discontinuously arrayed in a belt along the axis of the North American Cordillera, from northwest Mexico to British Columbia (for an overview see Coney, 1980).

Metamorphic core complexes are characterized by extensional mylonite zones, with younger on older faults in the allochthon, in contrast to mylonites and faults of the classic older on younger thrust and fold belts. Spectacular evidence for fluid activity accompanying deformation is present in the ubiquitous metasomatic chloritic breccias developed along detachment faults, in locally pervasive potassic and oxidative alteration of the allochthon, and sporadic Cu-Au mineralization associated with the chloritic breccias (Crittendon et al., 1980; Figure 8). Light stable isotope studies have been conducted to date on five metamorphic core complexes, the Picacho and Buckskins (Kerrich and Rehrig, 1985; Kerrich, 1987), Harcuvar (Roddy et al., 1988) complexes in central Arizona, the South Mountain complex of northern Arizona (Smith and Reynolds, 1985), and the Bitterroot Dome-Sapphire block detachment zone, western Montana (Kerrich and Hyndman, 1986).

*Mylonites*

The Picacho tectonic section features a pronounced upward trend in $\delta^{18}O$ values from 7.6 per mil in the infrastructure to 18.9 per mil in allochthonous Tertiary volcanic rocks, with an attendant decrease of estimated temperature from 550 to 150°C (Figure 20). Undeformed Oracle granite (1.6 Ga) of the lower tectonic plate retains near-magmatic mineral pair fractionations which correspond to isotopic temperatures of 500-600°C, that are typical of slowly cooled plutonic rocks (cf. Taylor, 1968). Mylonitic equivalents have whole rock $\delta^{18}O$ values in compliance with the granite protolith, but mineral-pair fractionations have been reset to 350-500°C, which probably records the ambient temperature of ductile deformation. Mylonites have been interpreted as forming either under closed-system conditions, or in the presence of fluids isotopically buffered by a rock reservoir of mylonitic equivalent. Comparable trends of $\delta^{18}O$ whole rock and isotopic temperature have been reported for mylonites in the Buckskins area. On the basis of H and O isotope systematics it is not possible to discriminate between magmatic and metamorphic fluids.

Ductile granite mylonites in the Bitterroot dome formed at ambient temperatures of 550 +/- 50°C in the presence of fluids having $\delta^{18}O$ values of 9 +/- 1 per mil (Figure 9), and likely of magmatic or metamorphic origin. Mylonites developed in metasedimentary rocks 25 km from the granite contact also retain isotopic concordance, which corresponds to temperatures of 450-500°C. Granite and metasediment mylonites formed under conditions of low water/rock ratio as indicated by the isotopic compliance of minerals and whole rocks in mylonites with their undeformed protolith.

*Chloritic breccias*

In the Picacho core complex chloritic breccias are shifted by ~+2 per mil relative to mylonites or fresh granites, and formed under open-system conditions at temperatures of 300 to 350°C, based on fluid inclusion, isotopic and mineralogical criteria. Hydrothermal fluids implicated in the Fe, Mg, Mn-metasomatism that produced the chloritic breccias have $\delta^{18}O$ values of 3 +/- 1 per mil and $\delta D$ values of -20 to -85 per mil (Figure 9). These were isotopically evolved meteoric waters or mixtures of evolved meteoric and metamorphic fluids. Veined and hydrothermally altered rocks of the middle plate record a progressive trend of $^{18}O$ enrichment up section, at diminishing temperatures (220-330°C), but the rocks exchanged with isotopically similar fluids as involved in the lower chloritic breccias. Granites and mylonite products in the Bitterroot lobe have tightly clustered $\delta^{18}O$ values, whereas feldspar and biotite have preferentially shifted to low $\delta^{18}O$ values and record disequilibrium fractionations. This effect is most pronounced in the chloritic breccias which are depleted by ~10 per mil relative to undisturbed mylonitic precursors; the breccias appear to have formed under conditions of lower temperature (250 to 370°C) and effective confining stress, acting as aquifers for the infiltration of $^{18}O$ depleted surface waters ($\delta^{18}O$ = -7 to -12 per mil). Criss and Taylor, (1983) demonstrated extensive O-and H-isotope exchange of the Idaho batholith and Bitterroot lobe with meteoric water, which occurred following tectonic removal of ~15 km of metasedimentary rocks in the Sapphire block cover (Kerrich and Hyndman, 1986).

At present, it is not clear as to whether brittle fracturing accompanied by meteoric water incursion was responsible for development of the chloritic breccias. Alternatively, the chloritic breccias could have been generated in the presence of more $^{18}O$ enriched fluids as was the case for the Picacho and Buckskins complexes, with subsequent isotopic overprinting by infiltration of meteoric water.

*Suprastructure*

Tertiary alkali basalts in the allochthonous upper plate of the Arizona core complexes are characterized by pervasive hydrothermal alteration to an assemblage of K-feldspar, calcite, Fe, Mn-oxides and anhydrite. $\delta^{18}O$ values of K-feldspar are highly erratic, spanning 6 to 21 per mil. $\delta^{18}O$ and $\delta^{13}C$ values of calcite are positively correlated with $\delta^{18}O$ values of K-feldspar, implying that some of the data spread can be accounted for in terms of temperature variation. Fluid inclusions in quartz indicate the presence of variably saline brines. Potassic alteration of the volcanic rocks appears to have been induced by formation brines stored within the upper plate, and derived largely from evolved meteoric water at temperatures of 100-250°C.

*Summary of fluid regimes*

The trends of increasing $\delta^{18}O$ values in conjunction with diminishing temperatures have been interpreted in terms of dual fluid regimes. At lower structural levels there is a switch from closed-system conditions in mylonites and granites to open system conditions in chloritic breccias, which reflect the expulsion of high temperature, crust-equilibrated and reduced fluids inducing Fe, Mg, Mn metasomatism: these fluids could have been magmatic, metamorphic or evolved formation brines or some combination of these reservoirs. In contrast, cool, oxidized thermal waters were present in the upper plate, likely of evolved

meteoric origin. Upper plate fluids locally penetrated down to the middle plate, mineralogically and isotopically overprinting the chloritic breccias. Thus the tectonic section in core complexes records an upward transition from high to low temperatures, low to enhanced water/rock ratios, ductile to brittle deformation, and the conjunction of deep aqueous reservoirs at lithostatic pressure with shallower surface reservoirs under hydrostatic conditions.

## TRANSPRESSIONAL TECTONIC REGIMES

Quartz and quartz carbonate veins are ubiquitous in metamorphic belts, notably in greenschist facies rocks. The majority of such veins are of cm to metre scale, and formed in response to pressure solution (Fyfe et al., 1978). Petrologic and isotopic studies have shown that such veins are generally compositionally and isotopically buffered by their wall rocks.

In contrast, giant quartz vein systems with vertical and lateral extents of kilometres are restricted to terrane boundary structures. Examples include the Destor-Porcupine and Kirkland Lake-Cadillac fault zones in the Archean Abitibi Southern Volcanic Zone, the Boulder-LeFroy fault in the Archean Norseman-Wiluna greenstone belt; the Jurassic Melones fault, foothills metamorphic belt, California; the Cretaceous-Tertiary Coast Range Megalineament, Juneau belt, Alaska; and the Insubric line, Southern Italian Alps of Cenozoic age. Quartz-carbonate veins, generally with more extensive carbonate alteration halos, are developed to variable intensities along those regional structures, but are most abundant in dilational jogs, duplexes and second or higher-order splays. The veins sporadically contain mesothermal Au-mineralization and, consequently, these veins in particular are the center of most studies. Kinematically the finite displacement vector may be high angle reverse fault, low angle thrust, oblique or transcurrent in orientation.

Irrespective of age, or the relative proportions of supracrustal lithologies, the vein systems and associated mesothermal precious metal deposits share a number of common features at a variety of scales. These include: (1) restriction to the brittle-ductile transition, (2) uniformly low fluid salinities (<3 wt.% NaCl equivalent), with $CO_2$ contents of 5-30 mole % (3) paragenesis dominated by quartz-carbonate, mica, chlorite, pyrite, scheelite and tourmaline, (4) vertical extents of up to 2 km and a lack of zoning, (5) restricted ranges of O, C, Sr and Pb isotopic compositions within vein systems, albeit with a geographic provinciality, (6) a late-kinematic timing, and (7) common association in space and time with shoshonitic magmatism (Boyle, 1979; Groves and Phillips, 1987; Wyman and Kerrich, 1988; Kerrich, 1989). These common features support the view that all such vein systems formed by a singular genetic process, irrespective of age.

The giant quartz vein systems in general, and mesothermal Au-Ag deposits specifically, have variously been attributed to lateral secretion, magmatic fluids, exhalative processes, structurally focused metamorphic outgassing, meteoric water circulation, or mantle degassing-granulitization. These metallogenic hypotheses have been critically reviewed by Kerrich, (1989), and only those aspects pertaining to the geochemistry of the fluids are dealt with here.

*Mantle CO2 - granulitization*

A number of workers have recently suggested that the fundamental process driving the ore-forming system of Archean Au-Ag vein deposits is granulitization, during late-Archean cratonization, induced by the advection of mantle $CO_2$ through lower crustal rocks (Colvine et al., 1988; Cameron, 1988). The granulitization model for gold deposits has been adapted from the model for granulite formation by influx of mantle-derived carbonic fluid, as proposed by Newton et al., (1980). The genetic scheme stems from three principal lines of evidence: (1) lithophile element (K, Rb, etc.) enrichments characteristic of the deposits (Bain, 1933; Kerrich and Fyfe, 1981; Davies et al., 1982) are considered to be directly complementary to the 'depletion' of these elements in granulites, (2) the existence of a secular conjunction of 'depleted' Archean granulites with Archean mesothermal Au deposits, and (3) calculated $\delta^{13}C$ values of the hydrothermal fluids carrying Au and LILE, many of which average -4 per mil, are equated with a juvenile mantle $CO_2$ reservoir ($-7\pm2$ per mil) that is presumed to have induced the granulitization and concurrent LILE depletion.

Tarney and Weaver, (1987) proposed that 'depleted' tonalitic Archean granulites are formed in a two-stage process, in which LILE are first extracted, followed by partial melting of the 'depleted' residue to generate tonalitic magmas. They specifically attribute these processes to an Archean subduction zone environment, under conditions of high geothermal gradient, thereby accounting for the observation that of all 'depleted' granulites, the majority are Archean. This scheme accounts for the codepletion of Th and U with K, Rb and Cs in granulites, whereas carbonic metamorphism cannot readily do so, owing to the fact that Th and U do not typically reside in minerals susceptible to dehydration by enhanced $CO_2$ activity (Tarney and Weaver, 1987).

It is well known that many post-Archean granulites, and some Archean examples, are not markedly depleted in the lithophile elements, especially those with sedimentary or volcanic protoliths (for a review see Rudnick et al., 1985), such that LILE-depletion is not an inevitable, or even common, result of granulite facies metamorphism. Accordingly, if many Archean granulites are born 'depleted' then the entire underpinning of the granulitization model for mesothermal quartz-gold veins is called into question, because there is no LILE complementarity between the two reservoirs.

Colvine et al., (1988) and Cameron, (1988) have selected those data from Archean mesothermal gold deposits where $\delta^{13}C \sim -3$ per mil as being directly equated with a juvenile $CO_2$ source, in support of the mantle $CO_2$-granulitization model. However, $\delta^{13}C$ values in this range are not uniquely diagnostic of a mantle carbon provenance (Kyser, 1986; Mattey, 1987; Matthews et al., 1987), nor are they representative of the larger population of $\delta^{13}C$ values in spanning $\sim$-11 to +1 per mil (Figure 10). Ohmoto and Rye, (1979) emphasized that $\delta^{13}C$ values of $\sim$-5 per mil are difficult to interpret in terms of the C source, precisely because mantle -$CO_2$, and average igneous, metamorphic and sedimentary carbon reservoirs all possess $\delta^{13}C$ values in this range. Alternatively, the spread of the entire population of $\delta^{13}C$ values has been attributed to mixtures of crustal and juvenile C (Golding and Wilson, 1983; Kerrich, 1983, 1987), where geographical variations in average $\delta^{13}C$ values may be related to crustal source rocks laterally heterogeneous with respect to oxidized ($^{13}C$ enriched) and reduced ($^{13}C$ depleted) carbon reservoirs.

There are several further weaknesses inherent to the mantle-$CO_2$ granulitization model for mesothermal vein systems. A key line of evidence used in support of the carbonic fluid metamorphism model of granulitization is the presence of primary-appearing, high density $CO_2$ fluid inclusions in granulite facies minerals (see Newton et al., 1980). Although in many instances these apparently primary inclusions have been shown to be related to a post-metamorphic uplift PT path (Lamb et al., 1987; Valley et al., 1988; Morrison and Valley, 1988). Vry et al., (1988) have shown that the $\delta^{13}C$ values of channel $CO_2$ in granulite facies cordierite is inconsistent with characteristic mantle $\delta^{13}C$ values, and that many granulites are internally buffered rather than being buffered by an external mantle $CO_2$ reservoir. O'Nions and Oxburgh (1988) have shown from $C/^3He$ measurements that the flux of mantle $CO_2$ in tectonically active areas is inadequate to produce granulites on a regional scale.

*Orthomagmatic models*

It is well known that many greenstone belt gold-quartz deposits are spatially associated with small intrusive stocks, rather than with granodiorite batholiths (Boyle, 1979; Kerrich and Fyfe, 1981; Colvine et al., 1988), but the significance of the association has remained elusive. Hattori, (1987) contended (1) that some of these stocks are analogs of I-type granites, (2) that the stocks have a primary oxidized character, and (3) that the gold is likely orthomagmatic. Alternatively, Fryer et al., (1979) proposed that the stocks were relatively brittle and acted as mechanically anisotropic bodies in layered volcanic-sedimentary sequences, thereby localizing hydraulic fracturing during post magmatic syntectonic fluid flow and mineralization: in this interpretation the spatial association with stocks is a consequence of contrasts in rheological properties.

Based on the sporadic presence of barite, anhydrite and hematite, Cameron and Hattori, (1987) suggested that the fluids in auriferous quartz veins were intrinsically oxidized, in accord with the apparently inferred oxidized nature of the felsic stocks. However, in many deposits the common occurrence of pyrite + arsenopyrite $\pm$ pyrrhotite assemblages, and locally graphite, constrains the fluid Eh to conditions at or below, the QFM buffer such that the model of Cameron and Hattori, (1987) is not general (Kishida and Kerrich, 1987). Moreover, transient immiscibility of $H_2$, $CH_4$, and $H_2S$ accompanying $CO_2$ immiscibility will act to oxidize residual fluid and $H_2S$, thereby increasing the $SO_4^{2-}$ activity, and in turn resulting in precipitation of sulphates and a concomitant shift $\delta^{34}S$ of sulphides to more negative values (Romberger, 1986; Spooner et al., 1987; Kerrich and Fyfe, 1988). Consequently, this oxidation, reflected in sporadic sulphate minerals, is largely a secondary feature evolved from a fluid originally close to QFM, rather than relating to a hypothetically oxidized magmatic reservoir.

A compelling case against any direct genetic link between mesothermal gold deposits and many of the spatially associated felsic intrusions can be made from geochronological considerations. Precise age determinations of felsic stocks and gold deposits in several areas of the Superior Province have shown that the latter significantly post-date synkinematic felsic magmatism, such that a direct orthomagmatic relationship of gold to the stocks can be ruled out (e.g. Colvine et al., 1988).

Another argument against purely magmatic hydrothermal fluids stems from the radiogenic isotope systematics of the vein systems. Most compositionally uniform magmas with rare exceptions (e.g. France-Lanord et al., 1988), possess uniform initial Sr, Pb and Nd isotope ratios. The vein systems, however, are characterized by highly variable Pb (Franklin et al., 1983; Dahl et al., 1987) and Sr-isotope initial ratios (Kerrich et al., 1987). The initial Sr isotope ratios of Archean trondhjemites specifically, are uniform and non-radiogenic (Peterman, 1979), in contrast to the dispersion of Sr isotope compositions of gold mineralization towards variably more radiogenic values (Figure 11). It is possible that magmatic fluids evolved by interaction with crustal rocks such that magmatic O, C, Sr and Pb isotope signatures were all modified; but in this case the orthomagmatic origin of the giant quartz vein systems and gold also becomes arguable.

*Meteoric water*

Nesbitt and coworkers proposed that the giant quartz vein systems that host mesothermal precious metal mineralization in the Cordillera formed in response to deep circulation of meteoric water along transcurrent faults under conditions of low water/rock ratio. They extended this model to Archean mesothermal veins (Nesbitt et al., 1986; Nesbitt, 1988; Nesbitt and Muehlenbachs, 1989).

The meteoric water model stems from the observation that $\delta$ D values of $H_2O$ obtained by decrepitating fluid inclusions in vein quartz from the Cordillera, apparently correlate with latitude, and therefore with $\delta$ D values of recent, and probably Tertiary, meteoric waters (Figure 12D). Two premises have been drawn from the conclusion that the hydrothermal fluids were of meteoric origin: (1) that mesothermal deposits are the deepest level expression of a vertical hydrothermal continuum, in which epithermal Au-Ag, Sb and Hg deposits represent successively higher levels in the system, and (2) that the fluid pressures approximated hydrostatic conditions in order to satisfy one of the criteria for Rayleigh convection (Nesbitt, 1988).

Pickthorn et al., (1987) critically evaluated the fluid inclusion data in mesothermal systems in the Cordillera. They concluded that the $\delta$ D results do not reflect primary fluid inclusion waters, derived from the hydrothermal fluids, but rather secondary inclusions that formed in the presence of surface meteoric waters infiltrating down the vein hosting structures, during uplift and late brittle fracturing. If this is correct, then a correlation of secondary fluid inclusion $\delta$ D values with latitude is expected, whereas the absence of a correlation is predicted for low water/rock ratios as should be required in models for mesothermal systems involving meteoric water (see below).

There are numerous independent lines of evidence that late infiltration of meteoric water does occur along the vein-hosting structures. Weir and Kerrick, (1984) reported late vuggy vein quartz precipitated from meteoric water in the Mother lode, and Kyser and O'Hanley, (1989) reported shifting to low $\delta$ D values of serpentine especially along faults in the Cassiar district, British Columbia. In the Canadian Shield, calcite and ferrodolomite paragenetically associated with vein quartz have been isotopically reset by later meteoric water, such that $\delta\,^{18}O$ values are variably out of equilibrium with quartz (Kerrich, 1989).

A problem also arises with the meteoric water model from the magnitude of $^{18}O$-shifts in the most northerly deposits. For example, at Cassiar, the total shift in $^{18}O$ from the meteoric water line would exceed 32 per mil, requiring extremely low fluid/rock ratios. Under the conditions of low fluid/rock ratio and moderate temperatures (300-500°C) required for pronounced $^{18}O$-shifts, meteoric waters evolve off the meteoric water line along j-shaped trajectories with significant positive $\delta$ D shifts. For a more realistic system, in which successive batches of water follow a specified temperature-water/rock pathway, still lower water/rock ratios are required because the H-isotope composition of the rock is progressively shifted to more negative $\delta$ D values. Accordingly, the model of Nesbitt et al., (1986) actually predicts the absence of a correlation between $\delta$ D values of the hydrothermal fluids and latitude.

Natural hydrothermal systems involving thermally driven convective circulation of fluids are characterized by a multiplicity of fluid pathways, each with a different temperature profile, and water/rock ratio, such that a variety of $\delta$ D and $\delta\,^{18}O$ shifts occur. This feature is observed in epithermal Au-Ag deposits in Nevada, where a wide span of $\delta$ D and $\delta\,^{18}O$ $H_2O$ values are observed (Figure 12). In contrast, mesothermal veins feature a restricted range of $\delta\,^{18}O$ quartz values and therefore uniform temperature and $\delta\,^{18}O$ $H_2O$ values. It is difficult to account for such uniformity in a convective system whereas fluid advection can readily do so.

There are also hydrodynamic problems inherent to the meteoric water model. Secor, (1985) showed that tensile fractures in crustal rocks require either a shallow depth of formation, or alternatively fluid pressures in excess of the hydrostatic pressure gradient (see also Etheridge et al., 1983). The depth of formation of mesothermal precious metal vein systems is not well constrained. Estimates based on fluid inclusion and mineral barometers converge on 1.5 to 3 kb, corresponding to depths of 6 to 12 km assuming lithostatic, or 15 to 30 km under hydrostatic, pressures (Ho, 1987; Weir and Kerrick, 1987; Walshe et al., 1988). At those depths tensile fractures that the veins occupy require suprahydrostatic fluid pressures, or where flat veins occur, supralithostatic pressures (Kerrich and Allison, 1978). Sibson et al., (1988) have shown that the reactivation of high angle normal faults, in which many of the quartz veins and mesothermal deposits occur, also requires suprahydrostatic fluid pressures. Meteoric water convection, on the other hand, occurs at close to hydrostatic fluid pressures.

Finally, the meteoric water model does not account for the observation that mesothermal vein systems are restricted to terrane boundary structures. The model would predict mesothermal veins in any large transcurrent fault with the appropriate attributes of fluid infiltration, thermal structure, and permeability: this is not the case.

In summary, the $\delta$ D values determined by Nesbitt et al., (1986) are consonant with other examples of meteoric water infiltration, such as the Alps (Figure 6), Canadian Shield, and terranes characterized by pronounced $\delta$ D-, but minor $\delta\,^{18}O$-, shifting (e.g. Kyser and Wilson, 1987), and where the isotopic overprinting is indicative of a secondary process rather than the primary hydrothermal regime.

*Metamorphic fluid systems*

The remaining hypothesis presented to account for the origin of giant quartz vein systems and lode gold deposits involves focused discharge of metamorphic fluids. Single-stage models for veins have been proposed by Fyfe and Henley, (1973), Kerrich and Fryer, (1979) and Fyfe and Kerrich (1986). Following arc-arc, or arc-continent transpressional collision, devapourization reactions of subcreted basalt and sediment generates significant fluid volumes during thermal rebound (Fyfe, 1973, 1974; Fyfe et al., 1978). Fluids are continuously released within the greenschist facies; and at the greenschist-amphibolite transition, which takes place at 450-500°C, about 5 wt.% structural water and volatiles including $CO_2$ are released from a hydrated rock of mafic composition. Large fluid volumes are also generated higher grades of metamorphism. Such fluids are channeled along terrane boundary faults and into second and higher order splays in the brittle-ductile transition where CO2 unmixing is promoted during pressure fluctuation. Metamorphic dehydration of lithologically complex sequences accounts for the complex H, O, C, Sr, and Pb isotope systematics of the veins.

*Acknowledgements*

We thank D. Kacher for wordprocessing the manuscript. R.K. and T.K.K. acknowledge NSERC operating grants, and the George McLeod endowment to the Dept. of Geological Sciences, University of Saskatchewan.

## References


Bain, G.W. (1933): Wall rock mineralization along Ontario gold deposits. Econ. Geol. 28, 705-745.

Boyle, R.W. (1979): The geochemistry of gold and its deposits. Geol. Surv. Can. Bull. 280, 584.

Cameron, E.M. (1988): Archean gold: relation to granulite formation and redox zoning in the crust. Geology 16, 109-112.

Cameron, E.M. and Hattori, K. (1987): Archean gold mineralization and oxidized hydrothermal fluids. Econ. Geol. 82, 1177-1191.

Colvine, A.C., Fyon, J.A., Heather, K.B., Marmont, S., Smith, P.M. and Troop, D.G. (1988): Archean lode gold deposits in Ontario: Ontario Geological Survey Miscellaneous Paper 139, 210 p.

Coney, P.J. (1980): Cordilleran metamorphic core complexes: an overview. Geol. Soc. Am. Mem. 153, 7-34.

Criss, R.E. and Taylor, H.P. Jr. (1983): An $^{18}O/^{16}O$ and D/H study of Tertiary hydrothermal systems in the southern half of the Idaho batholith. Geol. Soc. Am. Bull. 94, 640-663.

Crittenden, M.D., Jr., Coney, P.J. and Davis, G.H. (1980): Cordilleran metamorphic core complexes. Geol. Soc. Am. Mem. 153.

Dahl, N., McNaughton, N.J. and Groves, D.I. (1987): A lead-isotope study of sulphides associated with gold mineralization in selected deposits from the Archean Eastern Goldfield of Western Australia, in Ho, S.E. and Groves, D.I. (eds.), Recent advances in understanding Precambrian gold deposits: Geology Department and University Extension, University of Western Australia, Publication 11, 189-202.

Davies, J.F., Whitehead, R.E.S., Cameron, R.A. and Duff, D. (1982): Regional and local patterns of $CO_2$-K-Rb-As alteration: A guide to gold in the Timmins area, in Hodder, R.W. and Petruk, W. (eds), Geology of Canadian Gold Deposits: Canadian Institute of Mining and Metallurgy, Special Volume 24, 130-143.

Desmons, J. and O'Neil, J.R. (1978): Oxygen and hydrogen isotope compositions of eclogites and associated rocks from the E. Sesia Zone (W. Alps, Italy). Contrib. Mineral. Petrol. 67, 79-85.

Eslinger, E.V. and Savin, S.M. (1973): Oxygen isotope geothermometry of the burial metamorphic rocks of the Precambrian Belt Supergroup, Glacier National Park, Montana. Geol. Soc. Am. Bull. 84, 2549-2560.

Etheridge, M.A., Wall, V.J. and Vernon, R.H. (1983): The role of the fluid phase during regional metamorphism and deformation. J. Metamorph. Geol. 1, 205-226.

France-Lanard, C., Sheppard, S.M.F. and LeFort, P. (1988): Hydrogen and oxygen isotope variations in the High Himalaya peraluminous Manaslu leucogranite: Evidence for heterogeneous sedimentary source. Geochim. Cosmochim. Acta 52, 513-526.

Franklin, J.M., Roscoe, S.M., Loveridge, W.D. and Sangster, D.F. (1983): Lead isotope studies in Superior and Southern provinces. Geol. Surv. Can. Bull. 351, 60.

Frey, M., Hunziker, J.C., O'Neil, J.R. and Schwander, H.W. (1976): Equilibrium-disequilibrium relations in the Monte Rosa Granite, W. Alps: petrological, Rb-Sr and stable isotope data. Contrib. Mineral. Petrol. 55, 147-179.

Friedrichsen, H. and Morteani, G. (1979): Oxygen and hydrogen isotope studies in minerals from alpine fissures and their gneissic host rocks, W. Tauern Window (Austria). Contrib. Mineral. Petrol. 70, 149-152.

Fryer, B.J., Kerrich, R., Hutchinson, R.W., Pierce, M.G. and Rogers, D.G. (1979): Archaean precious metal hydrothermal systems. Dome Mine, Abitibi greenstone belt, I. Can. J. Earth Sci. 16, 421-439.

Fyfe, W.S. (1973): The granulite facies, partial melting and the Archean crust. Phil. Trans. Roy. Soc. Lond., Ser. A. 273, 457-461.

Fyfe, W.S. (1978): The evolution of the earth's crust: modern plate tectonics to ancient hot spot tectonics? Chem. Geol. 23, 89-114.

Fyfe, W.S. and Henley, R.W. (1973): Some thoughts on chemical transport processes with particular reference to gold. Mineral. Sci. Eng. 5, 295-303.

Fyfe, W.S. and Kerrich, R. (1985): Fluids and thrusting. Chem. Geol. 49, 353-362.

Fyfe, W.S., Price, N.J. and Thompson, A.B. (1978): Fluids in the Earth's Crust. Elsevier, Amsterdam, 383.

Gieskes, J.M. and Lawrence, J.R. (1976): Interstitial water studies, Leg 35. In: Initial Reports of the Deep Sea Drilling Project. 35, 407-424.

Golding, S.D. and Wilson, A.F. (1983): Geochemical and stable isotope studies of the No. 4 lode, Kalgoorlie, Western Australia. Econ. Geol. 78, 438-450.

Groves, D.I. and Phillips, G.N. (1987): The genesis and tectonic control on Archean gold deposits of the Western Australian Shield - a metamorphic replacement model: Ore Geology Reviews 2, 287-322.

Groves, D.I., Golding, S.D., Rock, N.M.S., Barley, M.E., and McNaughton, N.J. (1988): Archean carbon reservoirs and their relevance to the source of gold deposits. Nature 332: 254-257.

Hattori, K. (1987): Magnetic felsic intrusions associated with Canadian Archean gold deposits. Geology 15, 1107-1111.

Ho, S.E. (1987): Fluid inclusions: their potential as an exploration tool for Archean gold deposits, in Ho, S.E. and Groves, D.I. (eds), Recent Advances in Understanding Precambrian Gold Deposits. Extension Department, University of Western Australia, No. 11, 239-263.

Hoefs, J. (1978): Some peculiarities in the carbon isotope composition of "juvenile" carbons. Dept. Sci. Industrial Res. Bul. 220, 181-184.

Hoernes, S. and Friedrichsen, H. (1978): Oxygen and hydrogen isotope study of the polymetamorphic area of the Northern Otztal-Stubai Alps. Contrib. Mineral. Petrol. 67, 305-315.

Kerrich, R., (1983): Geochemistry of Gold Deposits in the Abitibi Greenstone Belt. Can. Inst. Min. Metall. Spec. Pap., 27: 75 pp.

Kerrich, R. (1988): Detachment zones of metamorphic core complexes: thermal, fluid and metasomatic regimes. Geol. Runds. 77/1, 157-181.

Kerrich, R. (1989): Shear zone hosted mesothermal gold deposits; a review of geochemical evidence on the sources of fluids and solutes, in Bursnall, J.T. (ed), Mineralization and Shear Zones. Geol. Assoc. Can. Short Course Notes 6.

Kerrich, R. and Allison, I. (1978): Vein geometry and hydrostatics during Yellowknife mineralization. Can. J. Earth Sci. 15, 1653-1660.

Kerrich, R. and Fryer, B.J. (1979): Archean precious-metal hydrothermal systems, Dome Mine, Abitibi Greenstone Belt: I. REE and oxygen isotope relations. Can. J. Earth Sci. 16, 440-458.

Kerrich, R. and Fyfe, W.S. (1981): The gold-carbonate association: source of $CO_2$ and $CO_2$-fixation reactions in Archean lode gold deposits. Chem. Geol. 33, 265-294.

Kerrich, R. and Hyndman, D. (1986): Thermal and fluid regimes in the Bitterroot Lobe-Sapphire Block detachment zone: evidence from $^{18}O/^{16}O$ and geological relations. Geol. Soc. Am. Bull. 97, 147-155.

Kerrich, R. and Kamineni, D.C. (1988): Chronology of fracture infilling at the Eye Dashwa Lakes granite, Atikokan, Ontario: evidence from stable isotopes and fluid inclusions. Contrib. Mineral. Petrol. 99, 430-445.

Kerrich, R. and Rehrig, W. (1987): Fluid motion associated with Tertiary mylonitization and detachment faulting: $^{18}O/^{16}O$ evidence form the Picacho metamorphic core complex, Arizona. Geology 15, 58-62.

Kerrich, R., Kamineni, D.C., Barre, D., Baldwin, D.K., McLarty, E. and Thivierge, R.H. (1987): Cyclic deformation and chemical transport in the Folsen lake Fault zone, East Bull Lake anorthosite-gabbro complex: evidence for seismic pumping? Appl. Geochem., 2, 103-126.

Kishida, A. and Kerrich, R. (1987): Hydrothermal alteration zoning and gold concentration at the Kerr-Addison Archean lode gold deposit, Kirkland Lake, Ontario. Econ. Geol., 82, 64-690.

Kyser, T.K. (1986): Stable isotope variations in the mantle. In Stable Isotopes in High Temperature Geological Processes, Valley, J.W., Taylor, H.P., Jr. and O'Neil, J.R. (editors), Mineral. Soc. Am., Rev. Mineral., vol. 16, Bookcrafters Inc., Michigan, 141-184.

Lamb, W.M., Valley, J.W. and Brown, P.E. (1987): Post-metamorphic $CO_2$-rich fluid inclusions in granulites. Contrib. Mineral. Petrol. 96, 485-495.

Lee, D.E., Friedman, I. and Gleason, J.D. (1985): Modification of $\delta D$ values in eastern Nevada granitoid rocks especially related to thrust faults. Contrib. Mineral. Petrol. 88, 299-298.

Mattey, D.P. (1987): Carbon isotopes in the mantle. Terra Cognita 7, 31-37.

Matthews, A., Fouillac, C., Hill, R., O'Nions, R.K. and Oxburgh, E.R. (1987): Mantle-derived volatiles in continental crust: the Massif Central of France. Earth Planet. Sci. Lett. 85, 117-128.

Morrison, J. and Valley, J.W. (1988): Post-granulite facies fluid infiltration in the Adirondack Mountains. Geology 16, 513-516.

Muehlenbachs, k. and Clayton, R.N. (1972a): Oxygen isotope studies of fresh and weathered submarine basalts. Can. J. Earth Sci. 9, 172-184.

Muehlenbachs, K. and Clayton, R.N., (1972b): Oxygen isotope geochemistry of submarine greenstones. Can. J. Earth Sci. 9, 471-478.

Nesbitt, B.E. (1988): Gold deposit continuum: A genetic model for lode Au mineralization in the continental crust. Geology 16, 1044-1048.

Nesbitt, B.E., Murowchick, J.B. and Muehlenbachs, K. (1986): Dual origin of lode deposits in the Canadian Cordillera. Geology 14, 506-509.

Nesbitt, B.E., and K. Muehlenbachs, Origins and movement of fluids during deformation and metamorphism in the Canadian Cordillera. Science 245, 733-736, 1989.

Newton, R.C., Smith, J.V. and Windley, B.F. (1980): Carbonic metamorphism, granulites and crustal growth. Nature 288, 45-50.

Ohmoto, H. and Rye, R.O. (1979): Isotopes of sulphur and carbon. In Geochemistry of Hydrothermal Ore Deposits, 2nd Edition, Barnes, H.L. (editor), John Wiley and Sons, New York, 509-562.

Oliver, J. (1986): Fluxes expelled tectonically from orogenic belts: their role in hydrocarbon migration and other geologic phenomena. Geology 14, 99-102.

O'Neil, J.R. and Taylor, H.P. Jr. (1967): The oxygen isotope and cation exchange chemistry of feldspars. Am. Mineral. 52, 1414-1437.

O'Nions, R.K. and Oxburgh, E.R. (1988): Helium volatiles fluxes and the development of continental crust. Earth Planet. Sci. Lett. 90, 331-347.

Peterman, Z.E. (1979): Strontium isotope geochemistry of late Archean to late Cretaceous tonalites and trondhjemites, in Trondhjemites, Dacites and Related Rocks. Elsevier, Amsterdam, 133-148.

Pickthorn, W.J., Goldfarb, R.J., and Leach, D.L., 1987. Comment on "Dual origins of lode gold deposits in the Canadian Cordillera". Geology, 15: 471-472.

Roddy, M.S., Reynolds, S.J., Smith, B.M., and Ruiz, J. (1988): K-metasomatism and detachment-related mineralization, Harcuvar Mountains, Arizona. Geol. Soc. Am. Bull. 100, 1627-1639.

Rudnick, R.L., McLennan, S.M. and Taylor, S.R. (1985): Large ion lithophile elements in rocks from high-pressure granulite facies terrains. Geochim. Cosmochim. Acta 49, 1645-1655.

Schidlowski, M. (1988): A 3,800-million-year isotopic record of life from carbon in sedimentary rocks. Nature (London) 333, 313-318.

Schoell, M. and Wellmer, F.W. (1981): Anomalous $^{13}C$ depletions in early Precambrian granites from Superior Province, Canada. Nature 290, 696-699.

Secor, D.T. (1965): Role of fluid pressure in jointing. Am. J. Sci. 263, 633-646.

Sibson, R.H., Robert, F. and Poulsen, H. (1988): High angle faults, fluid pressure cycling and mesothermal gold-quartz deposits. Geology 16, 551-555.

Smith, B.M. and Reynolds, S.J. (1985): Oxygen and hydrogen isotope study of mylonitization and detachment faulting, South Mountains, central Arizona. EOS Am. Geophys. Union Trans. 66, 1138.

Strauss, H. (1986): Carbon and sulfur isotopes in Precambrian sediments from the Canadian Shield. Geochim. Cosmochim. Acta 50, 2653-2662.

Tarney, J. and Weaver, B.L. (1987): Geochemistry of the Scourian complex: petrogenesis and tectonic models, in Park, R.G. and Tarney, J. (eds), Evolution of the Lewissian and Comparable High Grade Terrains. Geol. Soc. Lond. Spec. Publ. 27, 45-56.

Taylor, H.P., Jr. (1968): The oxygen isotope geochemistry of igneous rocks. Contrib. Mineral. Petrol. 19, 1-71.

Taylor, H.P., Jr. (1974): The application of oxygen and hydrogen isotope studies to problems of hydrothermal alteration and ore deposition. Econ. Geol. 69, 843-883.

Taylor, B.E. (1986): Magmatic volatiles: isotopic variation of C, H, and S. In Stable Isotopes in High Temperature Geological Processes, Valley, J.W., Taylor, H.P., Jr. and O'Neil, J.R. (eds), Rev. Mineral. 16, Mineral Soc. Amer., Bookcrafters Inc., Michigan, 185-225.

Thode, H.G., and Goodwin, A.M. (1983): Further sulphur and carbon isotope studies of late Archean iron formations of the Canadian Shield and the rise of sulphate reducing bacteria. Precambrian Research, 20, 337-356.

Valley, J.W., 1986. Stable isotope geochemistry of metamorphic rocks. In: J.W. Valley, H.P. Taylor, Jr., and J.R. O'Neil (editors), Stable Isotopes in High Temperature Geological Processes. Min. Soc. Amer., Revs. Mineral. 16: 445-490.

Vry, J., Brown, P.E., Valley, J.W. and Morrison, J. (1988): Constraints on granulite genesis from carbon isotope compositions of cordierite and graphite. Nature 332, 66-68.

Walshe, J.F., Kesler, S.E., Duff, D., and Cloke, P.L. (1988): Fluid inclusion geochemistry of high-grade vein-hosted gold ore at the Pamour Mine, Porcupine Camp, Ontario, Econ. Geol. 83, 1347-1368.

Weir, R.H. Jr. and Kerrick, D.M. (1984): Mineralogic and stable isotope relationships in gold-quartz veins in the southern Mother Lode, California. Geol. Soc. Amer. Programs with Abstr. 16, 688.

Weir, R.H., and Kerrick, D.M., 1987. Mineralogic, fluid inclusion and stable isotope studies of several gold mines in the Mother Lode, Tuolumne and Mariposa Counties, California. Econ. Geol., 82: 328-344.

Wickham, S.M. and Taylor, H.P., Jr. (1985): Stable isotope evidence of large-scale seawater infiltration in a regional metamorphic terrain; the Trois Seigneurs Massif, Pyrenees, France. Contrib. Mineral. Petrol. 91, 122-137.

Wilson, M.R., Kyser, T.K., Mehnert, H.H. and Hoeve, J. (1987): Changes in H-O-Ar isotope composition of clays during retrograde alteration. Geochim. Cosmochim. Acta 51, 869-878.

Wyman, D.A. and Kerrich, R. (1988): Archean lamprophyres, gold deposits and transcrustal structures: implications for greenstone belt gold metallogeny. Econ. Geol. 83, 454-459.

Yeh, H-W. (1980): D/H ratios and late-stage dehydration of shales during burial. Geochim. Cosmochim. Acta 44, 341-352.

Yeh, H-W. and Savin, S.M. (1977): Mechanism of burial of argillaceous sediments: 3. O-isotope evidence. Geol. Soc. Am. Bull. 88, 1321-1330.

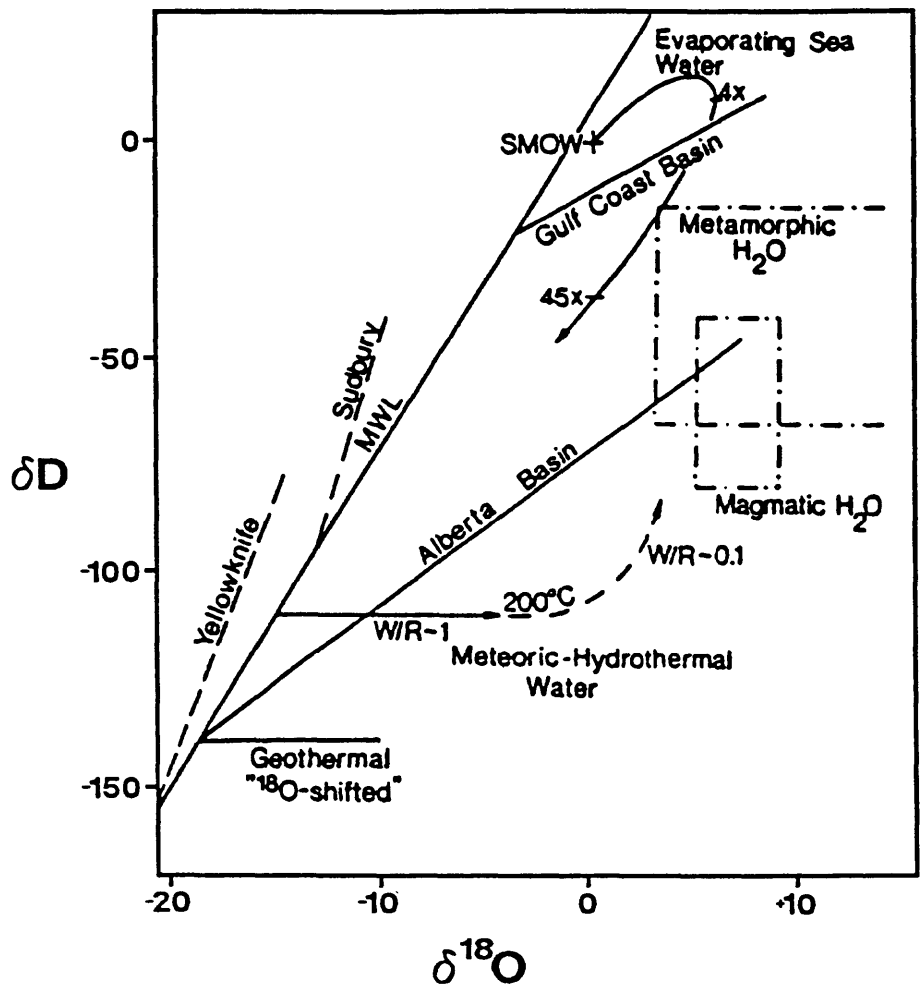



Figure 1. $\delta D$ and $\delta^{18}O$ values for various fluids associated with low temperature processes. Trajectory for evaporative seawater (SMOW) shows proposed increase in D and $^{18}O$ of the brine resulting from concentration in hydration spheres to about 75% evaporation (4x), and then decrease as evaporation proceeds. Highly evaporated brines have chemical and isotopic compositions that resemble mixing between evolved basinal brines such as those of the Gulf Coast and Alberta basins and modern meteoric water. The isotopic compositions of other waters from which minerals may form and fractionate isotopes include those along the meteoric water line (MWL), fluids in the Canadian Shield (Yellowknife and Sudbury), those associated with metamorphic rocks (Metamorphic $H_2O$) or magmas (Magmatic $H_2O$), geothermal fluids (Geothermal $^{18}O$-shifted), and meteoric waters that interact with igneous rocks in various water/rock (W/R) ratios (Meteoric-Hydrothermal Water).

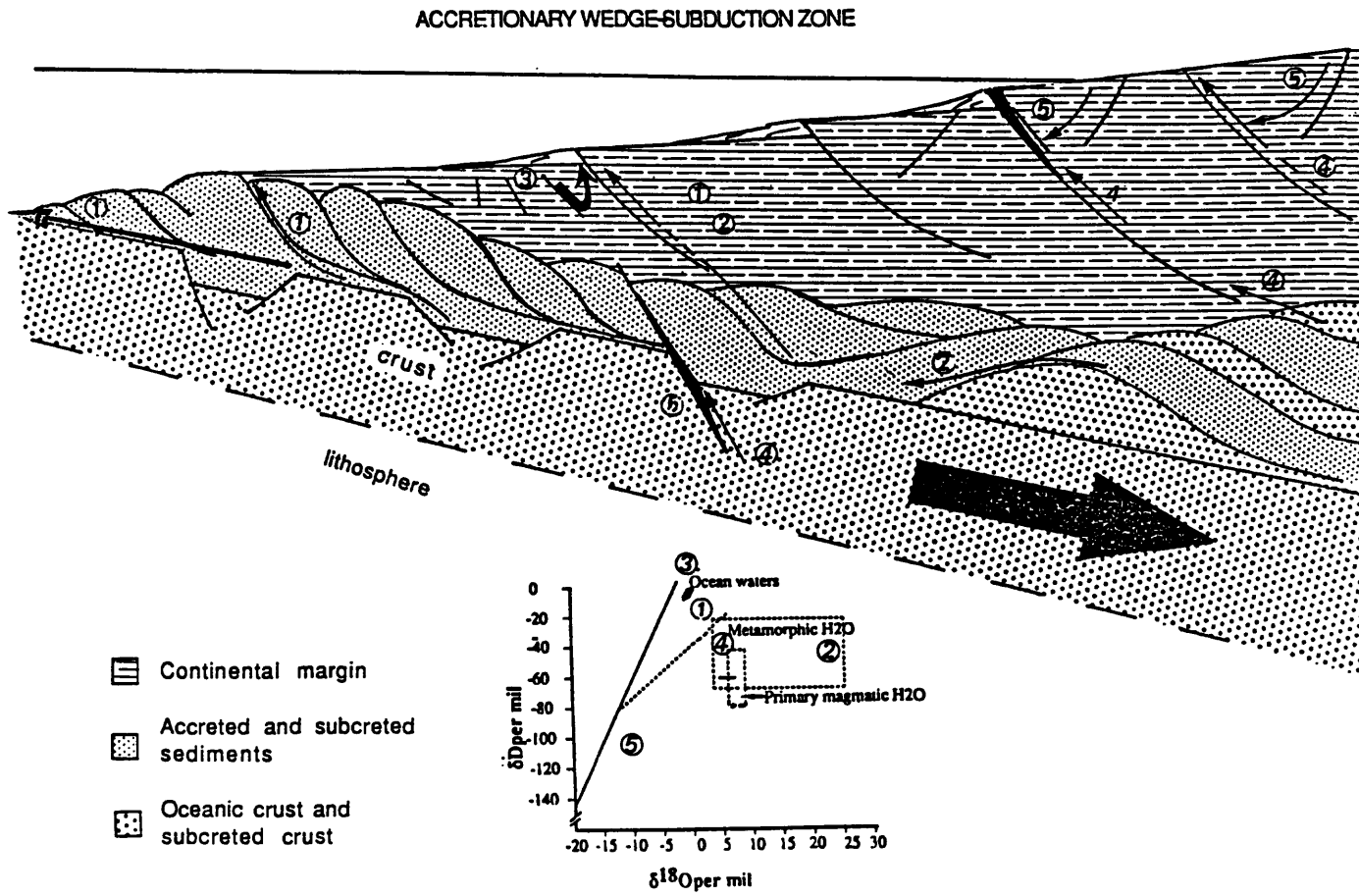



Figure 2. Schematic diagram illustrating possible fluid pathways in an accretionary prism and subduction zone. (1) Pore fluids dewatered from accreted sediments advect along decollement. (2) Fluids of dehydration from sediments. (3) Marine water convects through fractured continental margin. (4) Fluids dehydrated from oceanic crust and subcreted crust. (5) Meteoric water convects through fractured continental margin. (6) Serpentine diapirs extruded along faults. Fluid events not necessarily contemporaneous.

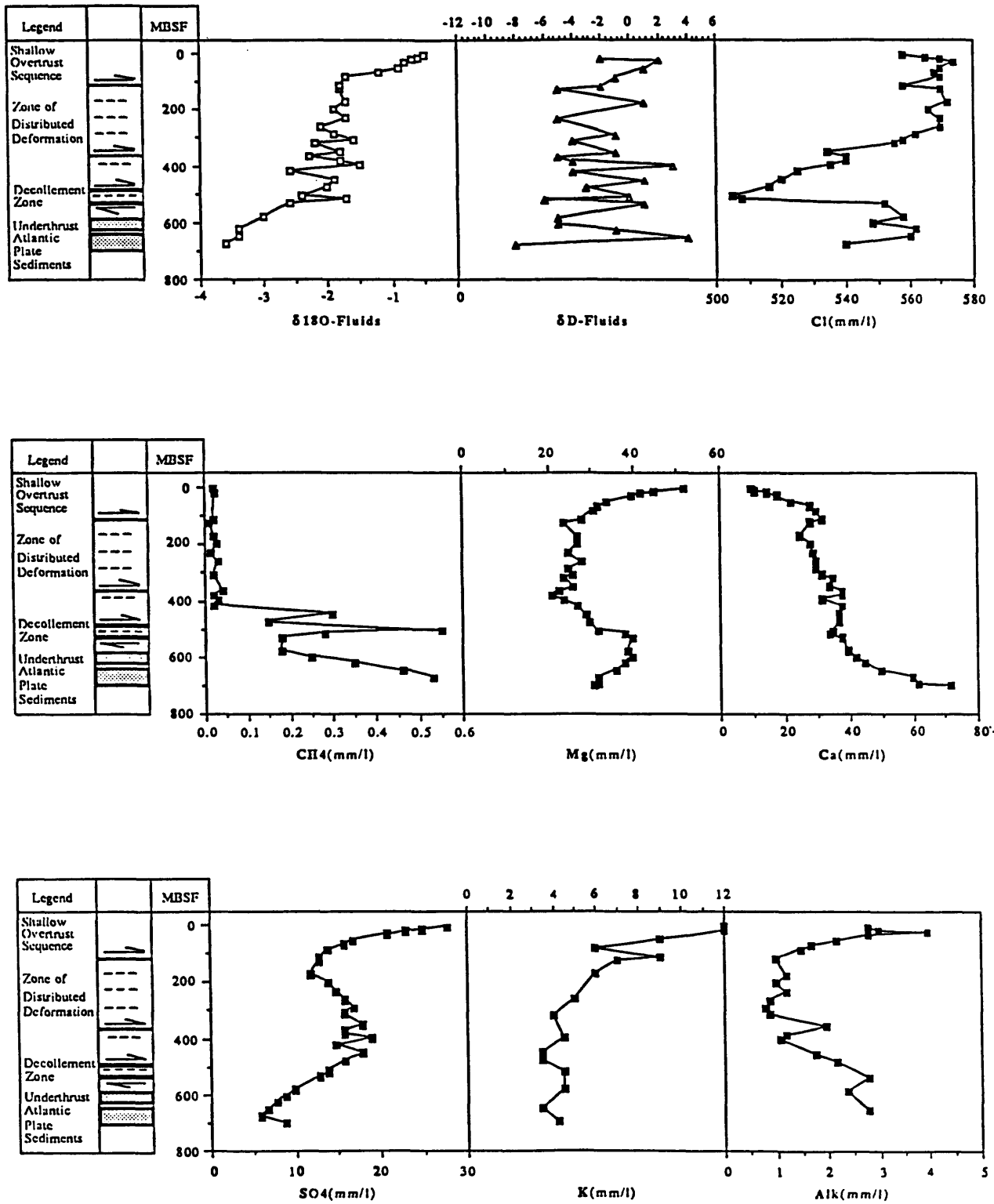



Figure 3. General stratigraphy and variation with depth in meters below sea floor (MBSF) of $\delta^{18}O$ and $\delta D$ values and chemical compositions of pore fluids in sediments from the Northern Barbados Accretionary Complex at site 671 of ODP Leg 110 (see Fig. C). Pore fluids near the decollement zone are relatively depleted in chloride but enriched in $^{18}O$, methane, $SO_4$, and perhaps deuterium. Data from Vrolijk et al. (in press).

**History of fluid advection, structural reactivation, and isotopic resetting in major structures, Archean Superior Province**

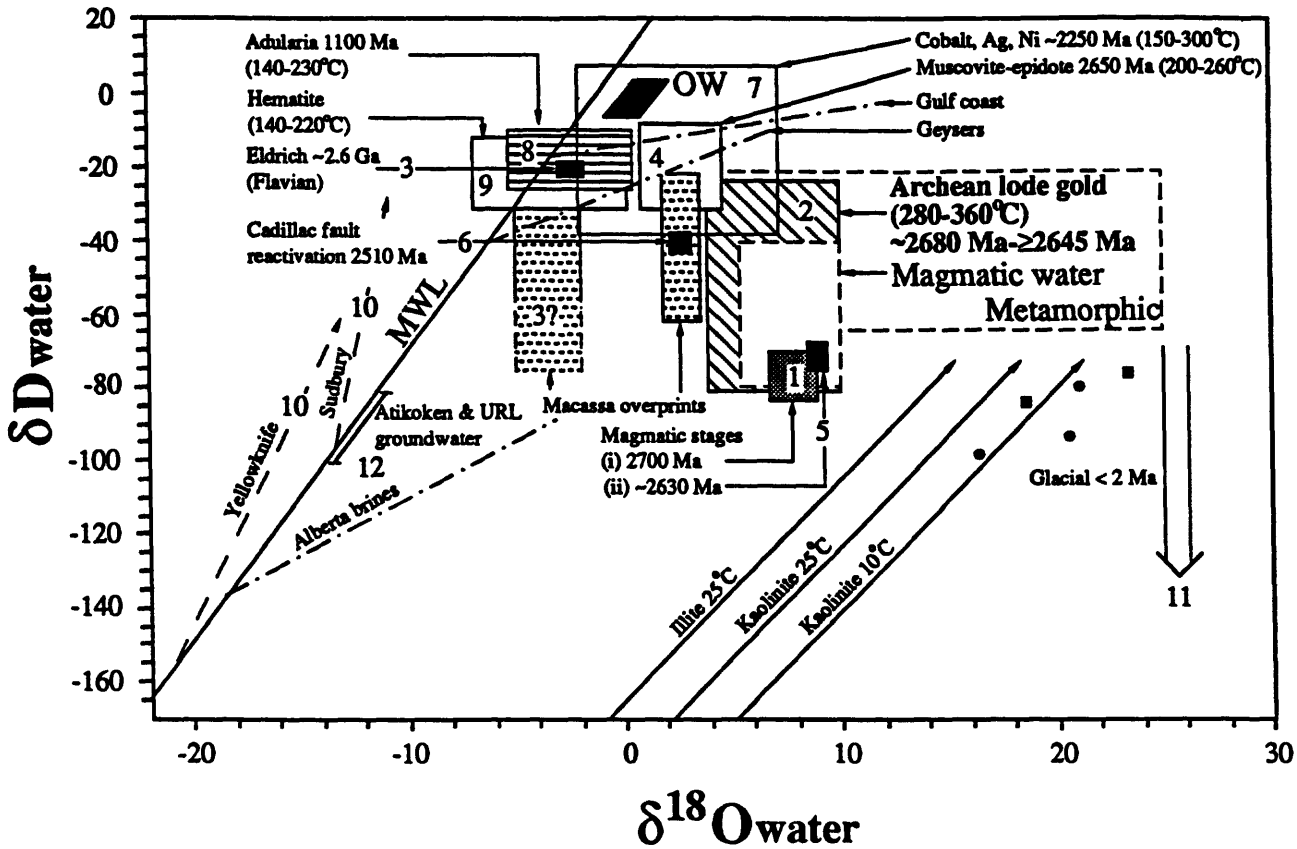



Figure 4. The hydrogen and oxygen isotope composition of specified brines in Precambrian basement. Yellowknife and Sudbury are trends for measured values of saline groundwaters, after Fritz and Frape, (1982) and Frape and Fritz, (1982). Also shown are calculated hydrogen and oxygen isotope compositions of fluids that infiltrated regional structures in the Archean Superior and Slave Provinces over ~2.7 Ga, selectively resetting primary stable and radiogenic isotope systems. Modified from Kerrich and Kamineni (1988) and Kerrich (1989). 1 - Peak metamorphic TTG magmatism at ~2700 m.y.; 2 - field for Archean auriferous ore-forming fluids 2680 m.y. to $\geq$ 2645 m.y.; 3 - Eldrich gold-deposits in the Flavrian Pluton, Abitibi SVZ, ~2700 m.y.; 4 - Superior Province wide muscovite-epidate veins; 5 - S-type granites ~2645 m.y. to 2611 m.y.; 6 - low-temperature fluid reactivation of Cadillac fault, west of Val d'Or at 2510 m.y.; 7 - ore-forming fluids for Proterozoic Ag, Co, Ni-sulpharsenide veins in Archean basement, Cobalt, Ontario; 8 - adularia bearing veins in Archean basement, 1100 m.y., 9 - hematite-dominated overprinting event; 10 - fluid reservoirs in Archean basement ~1500 m.y. to $\leq$ 2 m.y.; 11 - Quaternary ice age fluids infiltrating Archean structures inducing $\delta$D shifting (vertical arrow); 12 - modern groundwaters in Archean structures.

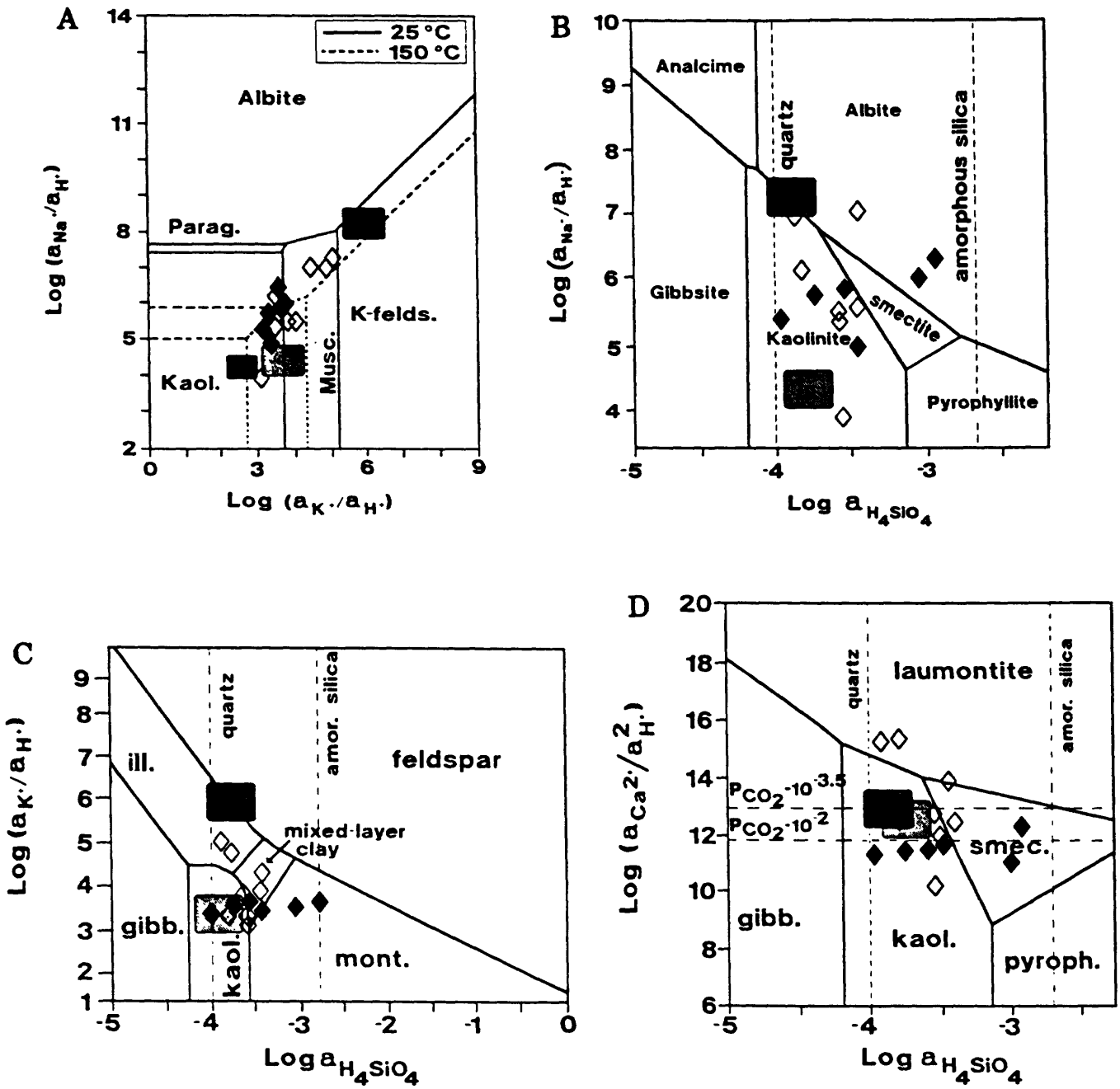



Figure 5. Equilibrium phase diagrams for the aqueous systems (a) $Na_2O$-$K_2O$, (b) $Na_2O$-$SiO_2$, (c) $K_2O$-$SIO_2$, and (d) CaO-$SiO_2$ and activity compositions of fluids from the Canadian Shield at Sudbury (open symbols) and Yellowknife (closed symbols). All systems at 25°C with the $K_2O$-$Na_2O$ system shown for both 25°C and 150°C. Also shown are quartz and amorphous silica saturation values and, for the system CaO-$SiO_2$, calcite saturation at $CO_2$ pressures of $10^{-3.5}$ (i.e. water in contact with atmosphere) and $10^{-2}$, and activity compositions for seawater (shaded area), river water (diagonal lines), and hydrothermal fluids from mid-ocean ridges (black area in $K_2O$-$Na_2O$ system).

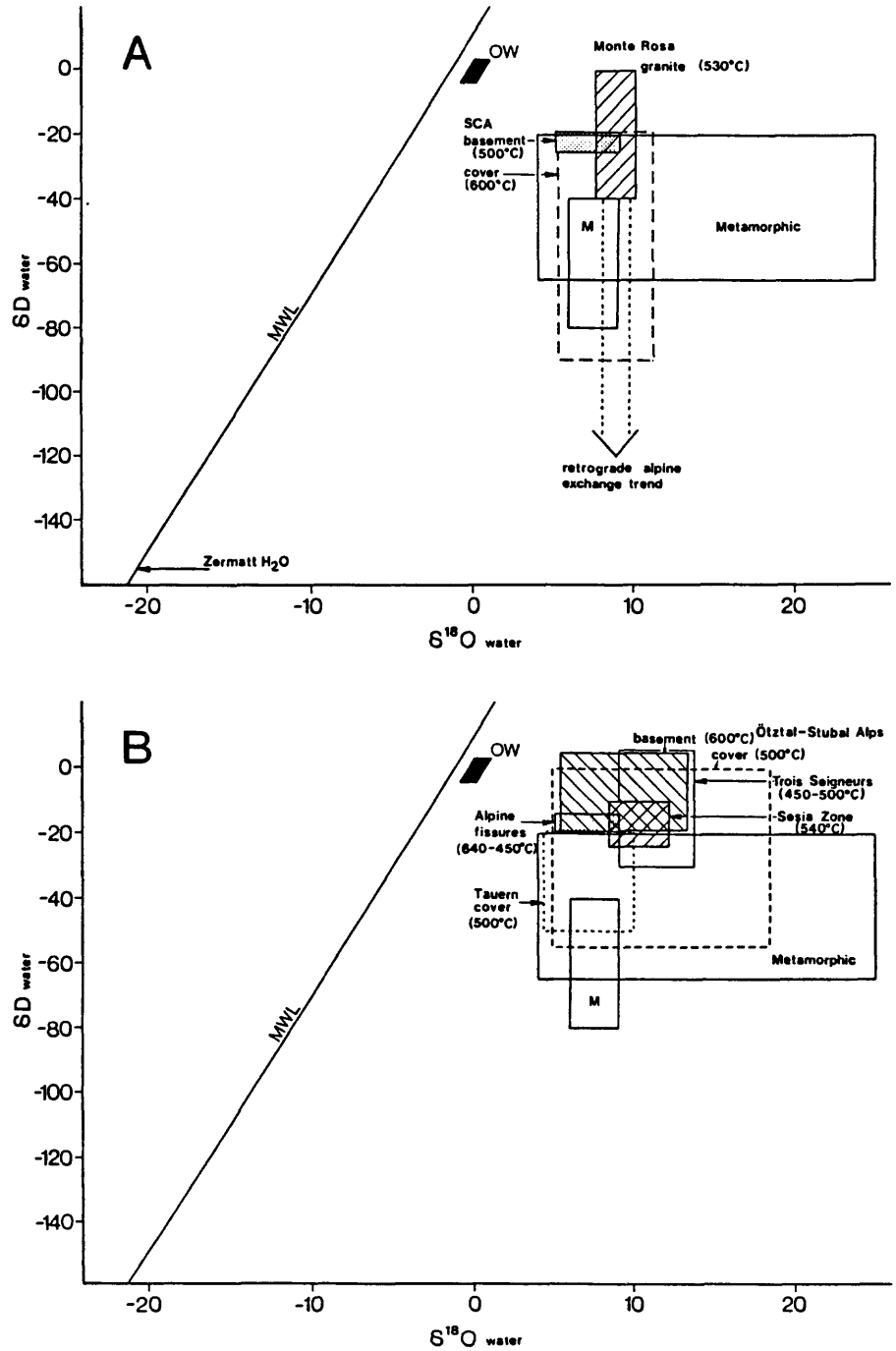



Figure 6. A. Calculated hydrogen and oxygen isotope compositions of waters in equilibrium with basement (solid outlines) and cover (dashed lines) tectonic units of the Alpine belt. Sources of data: Monte Rosa, Frey et al., (1976), Swiss Central Alps (SCA), Hoernes and Friedrichsen, (1978). MWL, OW, M and Metamorphic fields as in Figure 1. B. Calculated hydrogen and oxygen isotope compositions of waters in equilibrium with basement (solid outline) and cover (dashed outlines) tectonic units in the Alps. Otztal-Stubai Alps, Hoernes and Friedrichsen, (1978), Sezia Zone, Desmons and O'Neil, (1978), Alpine fissures and Tauern cover, Friedrichsen and Morteani, (1979). Also depicted is the field for Hercynian metamorphic basement of the Trois Seigneurs Massif, French Pyrenees (Wickham and Taylor, 1985). MWL, OW, M and Metamorphic fields as in Figure 1.

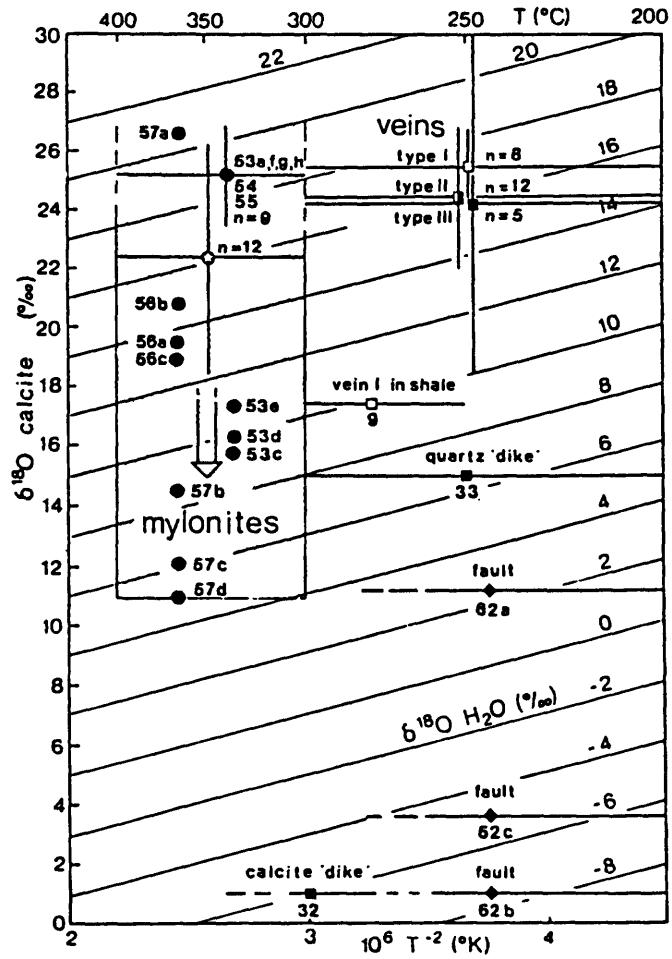



Figure 7. Measured individual $\delta^{18}O$ calcite values or ranges of values for selected calcites from specified generations of structures plotted versus the estimated temperature of formation. The family of diagonal lines represent $\delta^{18}O$ $H_2O$ in equilibrium with compatible values of $\delta^{18}O$ calcite and temperature according to the calcite-$H_2O$ equation of O'Neil et al., (1969). Horizontal lines are the estimated temperature range of a given structure, vertical lines represent two standard deviations of the mean $\delta^{18}O$ value, n is the number of data (after Burkhard and Kerrich, 1988).

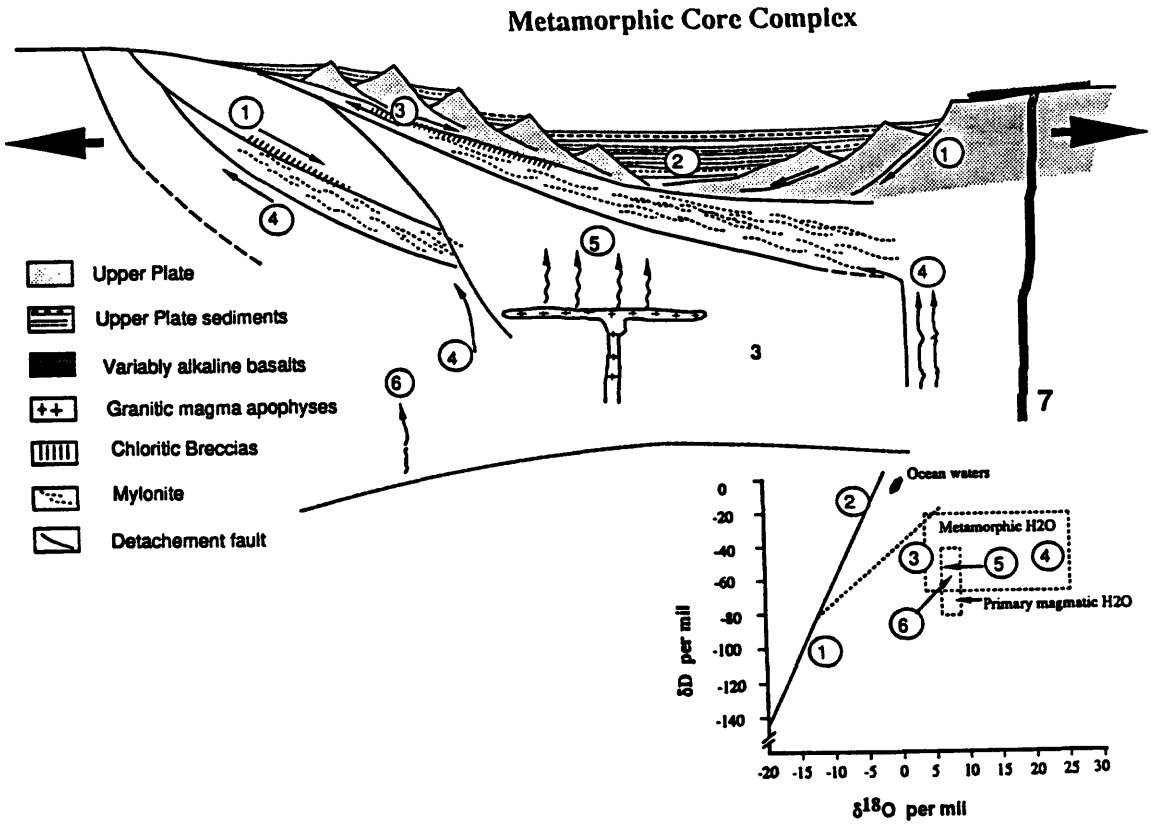



Figure 8. Schematic diagram showing the structure and possible fluid reservoirs involved in a metamorphic core complex, together with the oxygen and hydrogen isotope compositions of the fluids. (1) Meteoric water infiltrates down along listric normal; faults in upper plate, evolves by interaction with sediments in intermontane basin to hypersaline brines (2). Meteoric water also infiltrates into basement on left hand side, inducing variable H and O-isotopic overprint of basement and earlier formed chloritic breccias. (3). Zone of mixing between metamorphic fluids (4) advecting up detachment faults and meteoric water (1) infiltrating down. (4). Metamorphic fluids, generated by decompression and dehydration, and thermal envelope of crystallizing granites, advect up deep basement fractures into detachment faults. Such fluids may also include a component of extremely D and $^{18}O$-shifted meteoric water. (5). Fluids released from crystallizing syn-kinematic granites, generated by advection of heat into base off crust. (7). Mantle isotherms deformed upwards under stretched lithosphere: heat and basalts advected into base of lithosphere. (7). Variably alkaline basalts formed by low degrees of partial melting of decompressed asthenosphere.

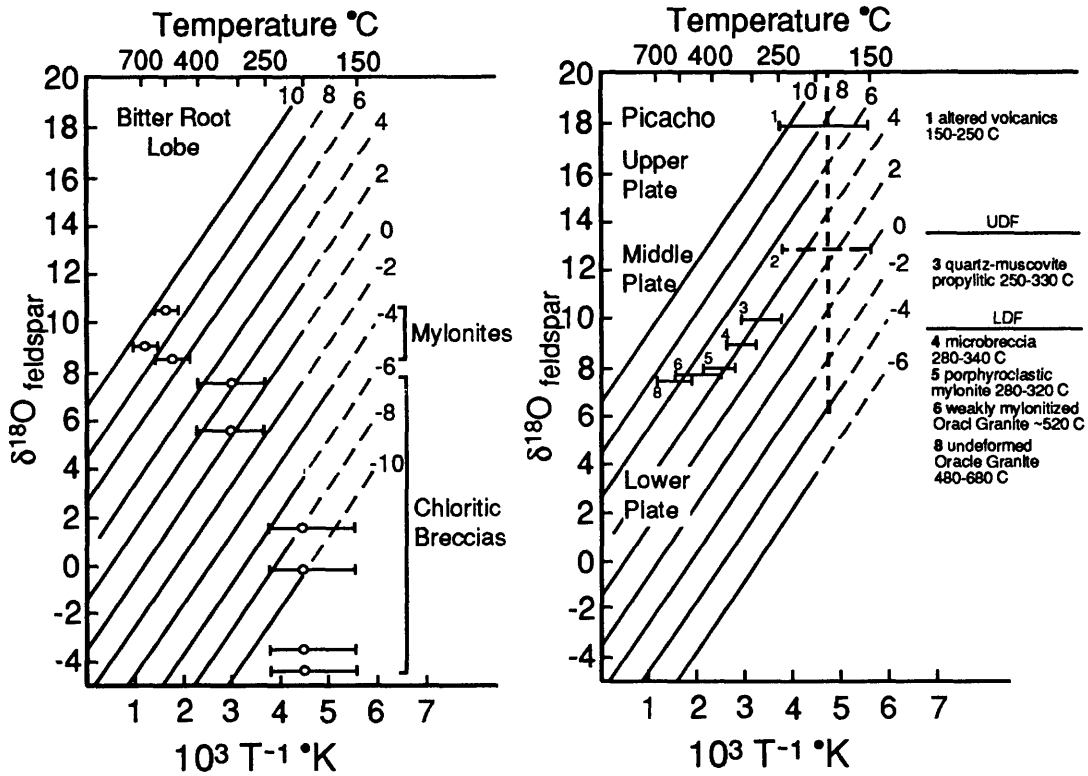



Figure 9. A. $\delta^{18}O$ values of feldspars (albite or K-feldspar) in granites, mylonitic derivatives and chloritic breccias from the Bitterroot lobe-Sapphire block detachment, with estimated equilibration temperatures. The family of diagonal lines are compatible values of $\delta^{18}O$ $H_2O$, temperature, and $\delta^{18}O$ feldspar, constructed from the feldspar-water calibration of O'Neil & Taylor, (1967). B. $\delta^{18}O$ values of feldspars (albitic plagioclase or K-feldspar) from granites, mylonites, chloritic breccias, and altered Tertiary alkali basalts, from the Buckskin Range metamorphic core complex, at Picacho Peak, with estimated temperatures of equilibration. The family of diagonal lines are as in A (modified from Kerrich, 1989).

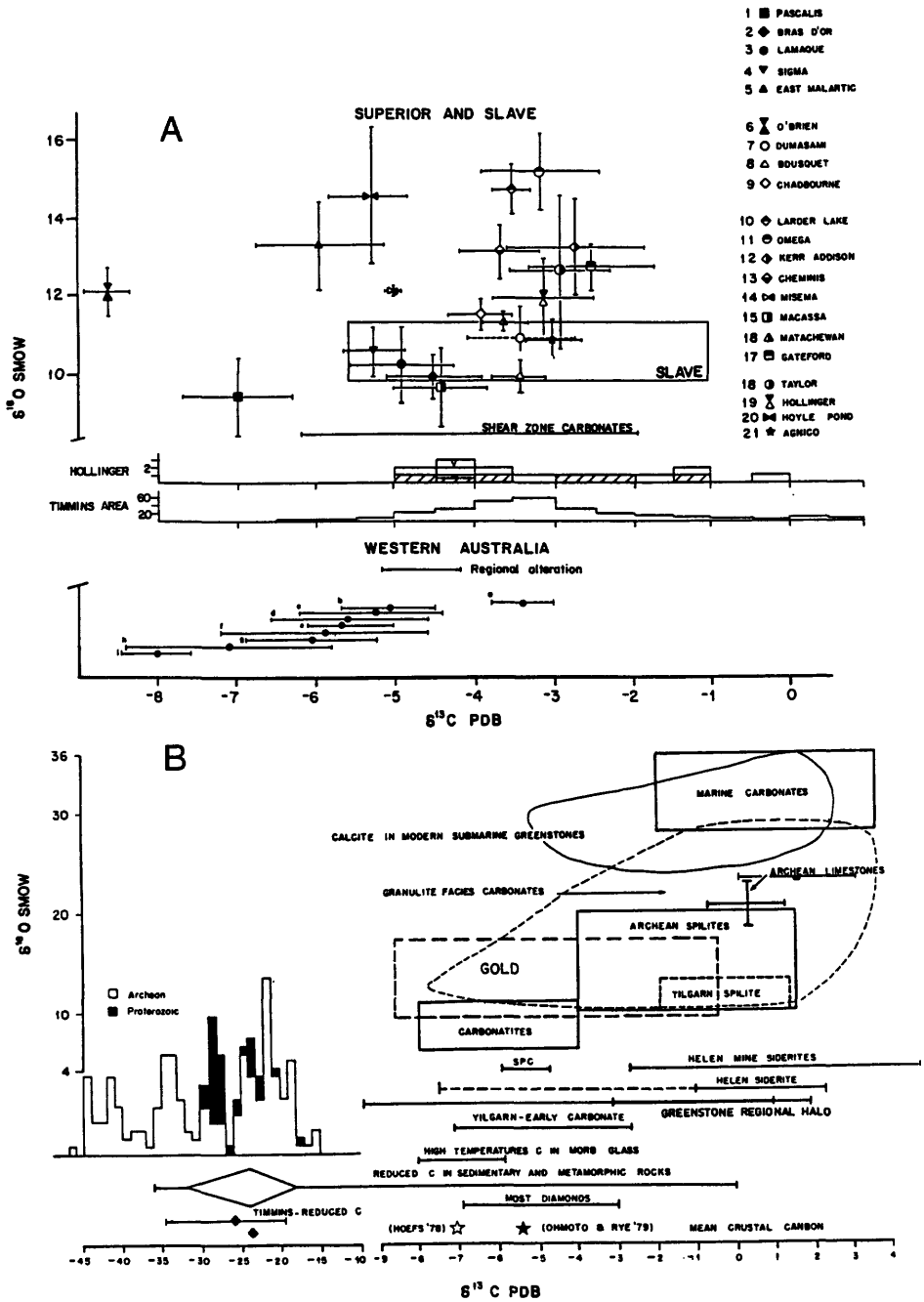



Figure 10. A. Average C and O isotope compositions of hydrothermal ferroan dolomite or calcite from specified mesothermal gold deposits in the Superior and Slave Provinces (after Kerrich et al., 1987). Western Australian data includes gold deposits from the Norseman-Wiluna belt (Groves et al., 1988). B. The O and/or C isotope fields of specified C reservoirs. Carbonatite field and granulite facies carbonates from Valley, (1986). Calcite in modern submarine greenstones from Muehlenbachs and Clayton, (1972a,b); high temperature C in MORB glass after Taylor, (1986). Field for marine carbonate, most diamonds and reduced C in metamorphic and sedimentary rocks, after Ohmoto and Rye, (1979). Estimates of average crustal C after Hoefs, (1978) and Ohmoto and Rye, (1979); reduced C in Superior Province from Schoell and Wellmer, (1981); Superior Province Carbonates (SPC) from Strauss, (1986); Archean organic C after Schidlowski, (1988); Yilgarn spilites and early carbonates from Groves et al., (1988). Helen mine siderites after Thode and Goodwin, (1983).

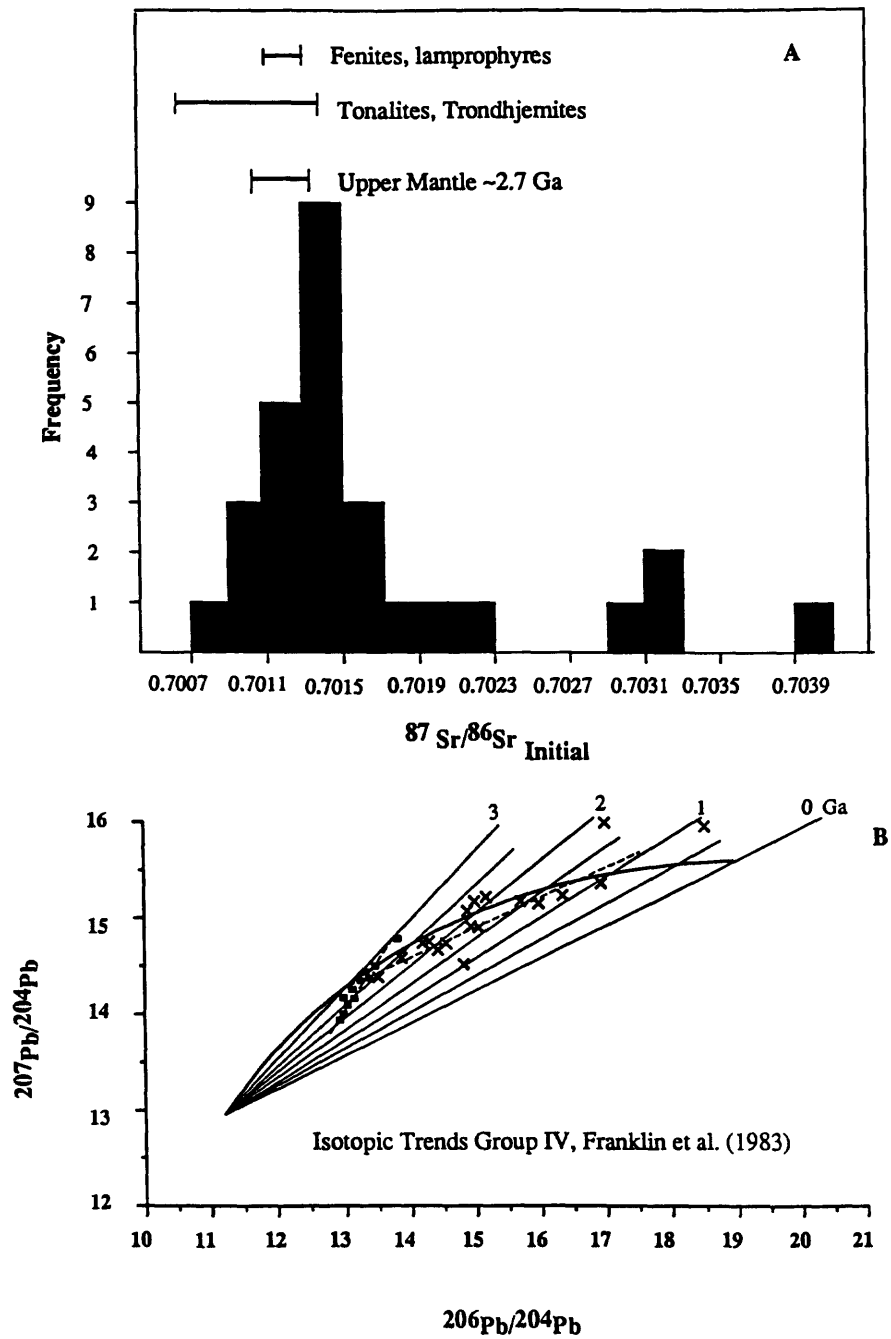



Figure 11. A. Contemporaneous $^{87}Sr/^{86}Sr$ of mineralized linear structures, compared with mantle, komatiites, and granodiorite batholiths, for the Abitibi Greenstone Belt at 2.7 Ga. B. Lead isotopic composition of massive sulphides (solid) and mineralized structures (crosses), after Franklin et al., (1983).

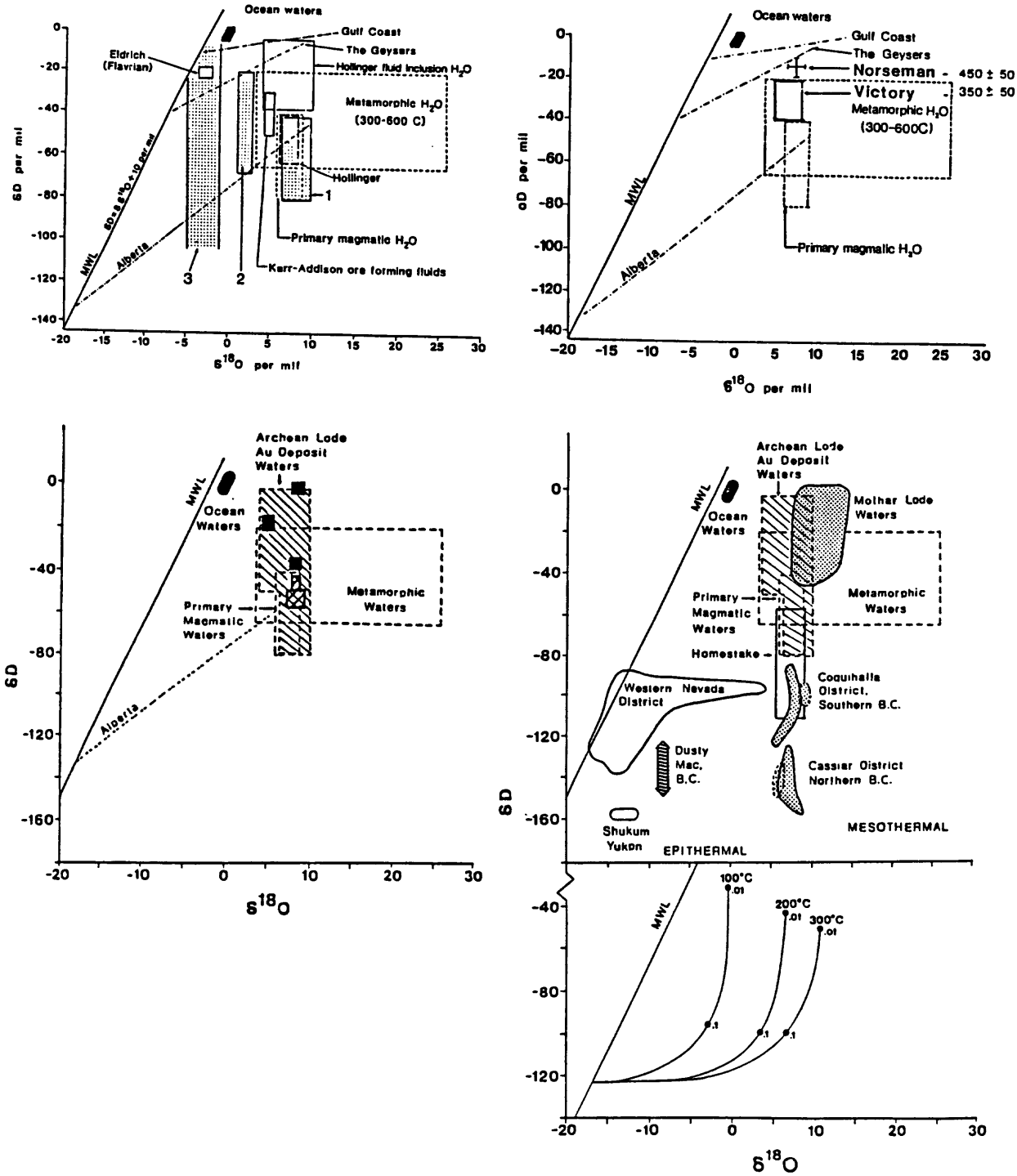



Figure 12. δD and $\delta^{18}O$ values of hydrothermal waters associated with giant quartz vein systems. A. Archean mesothermal deposits of Canada. B. W. Australia C. Proterozic deposits and D. Phanerozoic deposits. Sources in Kerrich (1989).

# 3: COUPLED MECHANICAL AND HYDROLOGICAL PROCESSES IN FAULTING

# Effects of Faulting Mode on Pore Fluid Stabilization of Slip

J. W. RUDNICKI
Departments of Civil Engineering, Mechanical Engineering and
Geological Sciences, Northwestern University, Evanston, IL 60208-3109

## ABSTRACT

This paper uses the model of *Rudnicki and Chen* [1988] to analyze alterations of frictional resistance due to changes in normal stress and pore fluid pressure on the slip surface. In particular, differences between normal and thrust faulting are emphasized. The model considers simple shear of a layer, and includes a slip-weakening relation between shear stress and slip, reduction of pore fluid pressure on the slip surface due to uplift and microcracking accompanying slip, and fluid flux to the slip surface in response to lowered pore fluid pressure. As pointed out by *Sibson* [1991], the compressive normal stress decreases with slip in the thrust geometry and, hence, decreases frictional resistance. Conversely, the normal stress increases with slip in the normal faulting and increases fricitonal resistance. The inclusion of pore pressure reduction due to dilatancy accompanying slip offsets, to some extent, these effects: The maximum pore fluid pressure reduction is greater for thrust faulting than for normal. The stabilizing effects vanish with increasing pore fluid compressibility and, hence, are limited by the ambient pore pressure prior to slip. This dependence of slip stability on the ambient pore fluid pressure suggests that the response will be sensitive to the local hydrologic structure near the fault.

## INTRODUCTION

*Sibson* [1991] has pointed out that simple models of earthquakes often assume implicitly that the frictional resistance to seismic slip is constant over the seismic cycle. He emphasizes that even if the frictional resistance is approximated as Coulomb, neglecting rate and state dependent effects, the resistance will typically change due to alterations in the normal stress and pore fluid pressure. He discusses qualitatively possible forms for these alterations based on the type of faulting (normal versus thrust) and the conditions of drainage near the fault zone.

The purpose of this note is to elaborate on the considerations of *Sibson* [1991] in the context of a model of unstable slip studied by *Rudnicki and Chen* [1988] (hereafter abbreviated *RC*). Although this model is quite simple, it does include some effects of fluid transfer between the slip surface and the surrounding material, more specifically, the competition between pore pressure reduction due to dilatancy accompanying slip and pore pressure increase due to fluid flow to the slip surface.

The model of *RC* was intended to analyze observations of *Martin* [1980] on the stabilization of failure of Westerly granite by pore fluid pressure effects. *RC* interpreted the observed stabilization as due to dilatant hardening accompanying postpeak slip on a fault formed near peak stress. Although there is considerable uncertainty in the values of some of the material and transport properties entering the analysis, the predictions for the critical pressure needed for stabilization and the trends with effective confining stress and nominal strain-rate agree reasonably well with those observed by *Martin* [1980].

The analysis of *RC* provides a simple model for unstable seismic slip and is used here to explore possible differences for normal versus thrust faulting. Although it might be expected

that the dilatancy observed to accompany slip in laboratory experiments would be suppressed *in situ* because of the larger normal stress, laboratory observations suggest otherwise. Specifically, Teufel's [1981] data show an increase in dilatancy with confining stress, presumably due to increased cracking near the slip surface caused by the indentation of hard asperity contacts. Furthermore, there appears to be ample, though not unequivocal, field evidence of open fissures near fault zones.

The next section reviews the model of *RC*. The following section discusses the differences in the predictions for normal versus thrust faults in the absence of pore fluid effects. Then the pore fluid effects are included. The concluding section discusses the results.

## MODEL

*RC* considered a slab of width $h$ loaded on one boundary by a prescribed shear displacement $U$, assumed to increase at a constant rate $V$, and a normal stress $\sigma$. This boundary is connected to a reservoir of pore fluid held at a constant pore pressure $p_{amb}$, where the subscript "*amb*" stands for "ambient". The other boundary is a frictional surface. This surface is assumed to have formed at peak stress and to sustain a shear stress $\tau$. The pore pressure on the slip surface is $p_{flt}$. Since the slab is infinite in extent, there is no dependence on distance along the slab. Furthermore, equilibrium requires that $\sigma$ and $\tau$ be independent of distance across the slab.

If the normal stress is constant, then the shear stress on the frictional surface decreases with slip $\delta$ from a peak value $\tau_p$ at zero slip to a residual value $\tau_r$ at a slip $\delta_0$:

$$\tau(\delta) = \tau_p - (\tau_p - \tau_r) g(\delta/\delta_0) \tag{1}$$

where $g(x)$ is a non-decreasing function with $g(0)=0$ and $g=1$ for $x \geq 1$. In the calculations *RC* use the following cubic function for $g$:

$$g(x) = -2x^3 + 3x^2 \tag{2}$$

When the effective normal stress is not constant, (1) is amended to

$$\tau(\delta) = \tau_p - (\tau_p - \tau_r) g(\delta/\delta_0) + m_0[\sigma - (p_{flt} - p_{amb})] \tag{3}$$

where $m_0$ is the coefficient of friction (assumed constant), $\sigma$ is the change in normal stress from its value at peak shear stress, and $p_{flt}$ is the pore fluid pressure on the slip surface. Since the pore fluid pressure is assumed to be uniform prior to slip, $p_{flt} = p_{amb}$ initially.

The material adjacent to the slip surface is assumed to be elastic. Thus, the shear stress is related to the elastic deformation of the layer by

$$\tau = \tau_p + G(U - \delta)/h \tag{4}$$

where $G$ is the elastic shear modulus. Equating (3) and (4) and rearranging in nondimensional form yields

$$T = Y - (2A/3)g(Y) + (\Sigma + P) \tag{5}$$

where $Y=\delta/\delta_0$, $T=Vt/\delta_0$, $P=m_0h(p_{amb}-p_{flt})/(G\delta_0)$, $\Sigma=m_0h\sigma/(G\delta_0)$, and $A=3(\tau_p - \tau_0)h/2G\delta_0$.

Dilatancy $\Delta$ is assumed to accompany slip. This dilatancy is not bulk volumetric strain [*Brace et al.*, 1966]. Rather it arises from the uplift in sliding over asperity contacts and microcracking induced near the sliding surfaces because of indentation of hard asperity contacts and local non-uniformities in slip. The following form is consistent with limited experimental data:

$$\Delta = \Delta_0 f(\delta/\delta_0) \tag{6}$$

where $\Delta_0$ is a constant. The function $f$ is assumed to be

$$f(x) = 2x - x^2 \tag{7}$$

in the calculations but, more generally, is an increasing function with $f(0) = 0$ and $f(1) = 1$. The form (6) applies for $\delta \leq \delta_0$. For $\delta > \delta_0$, dilatancy is assumed to saturate and $\Delta = \Delta_0$.

Deformation is coupled to the flow of pore fluid by assuming that a volume of fluid equal to that of the newly created void space, $\rho\Delta$ per unit area of slip surface, where $\rho$ is the mass density of pore fluid, is drawn into the slip surface. Thus, the rate of fluid mass influx $q$ is given by

$$q = \rho\dot{\Delta} + \rho\Delta p_{flt}/K_f \tag{8}$$

where the overdot denotes the time derivative and $K_f = \rho\, dp/d\rho$ is the bulk modulus of the pore fluid. The change in fluid mass on the slip surface must be balanced by flow from the surrounding material. This flow is assumed to be proportional to the difference between the pore pressure on the fault surface $p_{flt}$ and that in the reservoir $p_{amb}$:

$$q = \rho\kappa(p_{amb} - p_{flt})/h \tag{9}$$

where $\kappa$ is a permeability. This expression is exact (according to Darcy's law) only for steady state flow. Nevertheless, in view of the simplications that result, the uncertainty of material values, and other idealizations in the problem, the approximation seems reasonable. Equating (8) and (9), substituting (6), and rearranging in nondimensional form yields

$$2P = \epsilon f'(Y)\frac{dY}{dT} - \epsilon\alpha f(Y)\frac{dP}{dT} \tag{10}$$

where $\epsilon=V\,t_D\,/\delta_0$, $t_D=2\Delta_0 m_0 h^2/(\delta_0\kappa G)$, and $\alpha=G\delta_0/(m_0hK_f)$.

## NORMAL STRESS CHANGES

*RC* analyzed the equations outlined in the preceding section and presented numerical solutions for the case in which the normal stress is constant and, hence, can be eliminated from the formulation. In addition, they considered normal stress changes on the slip surface that simulate those in the axisymmetric compression test. Here, we review that formulation in the context of thrust faulting and, following *Sibson* [1991], contrast it with the effects of normal stress changes during normal faulting.

Let $\sigma_1 > \sigma_2 > \sigma_3$ be the principal stresses, positive in compression, and assume the slip surface is parallel to the direction of $\sigma_2$ and makes an angle $\theta$ with the direction of $\sigma_1$. The normal stress on the slip surface is

$$\sigma = \frac{1}{2}(\sigma_1 + \sigma_3) - \frac{1}{2}(\sigma_1 - \sigma_3)\cos 2\theta \tag{11}$$

and the shear stress $\tau_s$ is

$$\tau_s = \frac{1}{2}(\sigma_1 - \sigma_3)\sin 2\theta \tag{12}$$

For thrust faulting, the minimum principal stress is vertical and constant, $\sigma_3 = \sigma_v$. The maximum principal stress is horizontal, $\sigma_1 = \sigma_h$ and slip is initiated by an increase in $\sigma_1$. Therefore changes in normal stress $\Delta\sigma$ are related to changes in shear stress $\Delta\tau$ by

$$\Delta\sigma = \Delta\tau(1 - \cos 2\theta)/\sin 2\theta \tag{13}$$

As emphasized by *Sibson* [1991], changes in shear stress are accompanied by changes in normal stress of the same sign. Consequently, a drop in shear stress with slip causes a drop in the frictional resistance to slip because of the decrease in normal stress. Substituting (13) into (3) (with $p_{flt} = p_{amb}$) and rearranging yields an equation of the same form with the difference in peak and residual stress $\tau_p - \tau_r$ replaced by $(\tau_p - \tau_r)\gamma_t^{-1}$, where

$$\gamma_t = 1 - m_0(1 - \cos 2\theta)/\sin 2\theta \tag{14}$$

and the subscript "t" denotes the value for thrust faulting. Thus, the normal stress decrease increases the effective shear stress drop. For $m_0 = 0.6$ and $\theta = 30°$, typical values, $\gamma_t = 0.65$ and the effective shear stress drop is increased by more than 50%.

For normal faulting, the maximum principal stress is vertical and constant, $\sigma_1 = \sigma_v$, and the minimum principal stress is horizontal, $\sigma_3 = \sigma_h$. Slip is initiated by a decrease in horizontal stress (typically accompanying extension). Therefore changes in normal stress $\Delta\sigma$ are related to changes in shear stress $\Delta\tau$ by

$$\Delta\sigma = -\Delta\tau(1 + \cos 2\theta)/\sin 2\theta \tag{15}$$

In constrast to thrust faulting, the changes in normal stress are opposite in sign to the changes in shear stress. Substituting (15) into (3) and rearranging again yields an equation of the same

form with the difference in peak and residual stress $\tau_p - \tau_r$ replaced by $(\tau_p - \tau_r)\gamma_n^{-1}$, where

$$\gamma_n = 1 + m_0(1 + \cos 2\theta)/\sin 2\theta \tag{16}$$

and the subscript "n" denotes the value for thrust faulting. For $m_0 = 0.6$ and $\theta = 30°$, typical values, $\gamma_n = 2.04$ and the effective shear stress drop is increased by a factor of about two.

When the normal stress changes for either normal or thrust faulting are used to eliminate $\Sigma$ from (5), the equation can be rearranged to yield

$$T = Y - (2A/3\gamma)g(Y) + P/\gamma \tag{17}$$

If pore fluid changes are neglected (drained deformation, $P = 0$), then the ratio of an increment of slip to an increment of boundary displacement is

$$\frac{d\delta}{dU} = \frac{1}{1 - (2A/3\gamma)g'(Y)} \tag{18}$$

where the prime denotes the derivative. Slip becomes unstable, corresponding to an unbounded derivative in (18), when the denominator vanishes. This instability has the interpretation that the inertial terms that have been neglected in the formulation become significant and slip is seismic.

For g(Y) given by (2) instability occurs at $Y = Y_*$, where

$$Y_* = \frac{1}{2}[1 - (1 - \gamma/A)^{1/2}] \tag{19}$$

A second value of $Y$ which satisfies (17) with $P = 0$ for the same value of $T$ is

$$Y_1 = (1/2) + (1 - \gamma/A)^{1/2} \tag{20}$$

The condition for instability is met whenever $A > \gamma$. Therefore, for a given material $\tau$ versus $\delta$ curve, layer height, and modulus, instability is more likely for normal stress changes typical of thrust faulting. If instability does occur, it occurs at a smaller value of $\delta$ and with a larger stress drop for thrust faulting than for normal faulting. (Seismological studies seem to find no systematic differences in the stress drops for normal and thrust faulting. However, stress drop determinations are not precise and the differences suggested here, factors of four, may not be large enough to detect easily among other variations. Also, the stress drop here, the difference between peak and residual stress, is not identical to a value determined during dynamic slip.)

## Pore Pressure Stabilization

*RC* showed that when pore pressure reduction due to dilatancy is included the instability described in the preceding section is prevented unless $A$ or $\alpha$ is large. Large $A$ corresponds to large stress drop or a compliant elastic surroundings; large $\alpha$ corresponds to a compressible pore fluid. Because $\alpha$ is proportional to the pore fluid compressibility $(1/K_f)$, its value depends on the absolute pore pressure on the slip surface and, hence, on the ambient value prior to reduction

caused by dilatancy: Since the absolute pore pressure cannot be less than zero, the total reduction cannot exceed the ambient value. If the pore pressure is reduced sufficiently, the bulk modulus of the pore fluid $K_f$ decreases and the stabilization is diminished. *RC* argue that for a pore fluid of liquid water and shallow crustal conditions, the reduction in $K_f$ occurs so abruptly that it can be taken as constant down to a small critical value of pressure. Temperature increases due to frictional heating or the presence of dissolved gases may, however, alter the behavior of $K_f$.

In this section, we examine the differences in pore fluid stabilization for thrust and normal faulting. Determining the response including pore pressure changes due to dilatancy requires that (10)) and (17) be solved simultaneously. These two equations can be combined by using (17) to eliminate $P$. The result is

$$\epsilon \frac{dY}{dT} = \frac{2[T-Y+(2A/3\gamma)g(Y)]+\epsilon\alpha f(Y)}{f'(y)/\gamma + \alpha f(Y)[1-(2A/3\gamma)g'(y)]} \tag{21}$$

Although the solution of (21) requires numerical integration, it is evident that the slip rate (left hand side) remains bounded and, hence, instability does not occur unless the denominator on the right hand side vanishes. This denominator vanishes only if the undrained response (corresponding to no fluid flux into the dilating slip zone) is unstable.

*RC* discuss in detail the case $\gamma=1$ (corresponding to no normal stress changes). In particular, their Figure 4 plots the locus of values of $A$ and $\alpha$ which separates regions in which the denominator has a root in $0< Y\leq 1$ from those which it does not. For $\alpha = 0$, corresponding to an incompressible pore fluid, the response is always stable (denominator has no root) regardless of the value of $A$. As $\alpha$ becomes larger, the transition from stable to unstable response occurs at a finite value of $A$ that decreases with increasing $\alpha$. In the limit as $\alpha$ approaches $\infty$ (highly compressible pore fluid), the stabilizing effect vanishes and instability occurs for the same value of $A$ as for drained response, i.e., $A = \gamma$.

The results of *RC* can be generalized to include normal stress changes ($\gamma \neq 1$) by making the following replacements in the graph in their Figure 4: replace $A$ by $A/\gamma$ on the ordinate and $\alpha$ by $\alpha\gamma$ on the abscissa. The effects of these alterations on the transition from stable to unstable response can be assessed by a numerical example. For liquid water, the pore fluid bulk modulus $K_f$ is about 2.2 GPa. Using this value with $G$ = 30 GPa, $\delta_0$ = 0.2 mm, $m_0$ = 0.6 and $h$ = 2 cm, values *RC* suggest are appropriate for modelling *Martin's* [1980] laboratory experiments, yields $\alpha$ = 0.23. For this value of $\alpha$ and $\gamma=1$ (normal stress constant), the minimum value of A needed to cause instability of the undrained response is 4.5. For $\gamma=0.65$, corresponding to thrust faulting, the critical value of $A$ is 3.85. For $\gamma=2.0$, corresponding to normal faulting, the critical value of $A$ is 5.7.

Uncertainty about the value of $\delta_0$ and the one dimensional nature of the model make it difficult to estimate accurately the value for $\alpha$ *in situ*. It does seem likely that the ratio $\delta_0/h$ will be smaller *in situ*, resulting in a smaller value of $\alpha$ but this could be offset by a reduced value of $K_f$. For the purposes of illustration, we take $\alpha$ = 0.1. For constant normal stress ($\gamma=1$), thrust faulting ($\gamma=0.65$), and normal faulting ($\gamma=2.0$), the minimum values of A needed to cause instability of the undrained response are 8.1, 7.5, and 9.6, respectively.

Recall that for drained response ($P=0$), the minimum value of $A$ needed to cause instability is $\gamma$. For normal versus thrust faulting, $\gamma$, and the critical value of $A$, varies by a factor of about 3, but the critical value of $A$ for undrained response varies by only a factor of

1.3 for $\alpha = 0.1$ and 1.5 for $\alpha = 0.23$. This indicates that the effects of normal stress changes on stability can be offset by the reductions in pore pressure. For example, the destabilizing reduction in normal stress that accompanies slip in the thrust fault geometry is mitigated by a larger reduction in the pore fluid pressure. This feature may be another factor contributing to the absence of observed systematic differences in the stress drops of normal versus thrust earthquakes.

*RC* note that the main features of the solution, in particular, the maximum pore pressure reduction, can be determined by an asymptotic analysis. This analysis is motivated by the small value of $\epsilon$; *RC* estimate that the range $10^{-3}$ to $10^{-6}$ is representative of both field and laboratory conditions. A result of this analysis is that the nondimensional pore pressure reduction, $P$, is closely approximated by

$$P = (4/3)A(Y - Y_*)^2(Y_1 - Y) \tag{22}$$

The maximum pressure reduction is

$$P^{max} = (2/3)A(1 - \gamma/A)^{3/2} \tag{23}$$

Taking $\gamma = 0.65$ for thrust faulting and $\gamma = 2.0$ for normal faulting, the ratio of the maximum pressure reduction for normal faulting to that for thrust faulting varies from 0.28 to 0.79 for $A$ ranging from 3 to 10. Thus, the pressure reduction for normal faulting is less than for thrust faulting. Because the minimum pressure on the slip surface must be positive, the ambient pressure must exceed the pressure reduction in order for stabilization.

## DISCUSSION

As pointed out by *Sibson* [1991], the normal stress changes in the thrust and normal faulting geometries change the frictional resistance according to the simple Coulomb criterion. In particular, the normal stress decreases during slip in the thrust geometry, decreasing resistance; conversely, the normal stress increases during slip in the normal fault geometry, increasing resistance. The estimates of *Sibson* [1991] and those here suggest that the magnitude of the effect can be significant. For a fault that makes a 30° angle with the maximum compressive stress and a friction coefficient of 0.6, the *RC* model adopted here predicts that the effective shear stress drop from peak to residual is increased by a factor of 1.5 for thrust geometry and decreased by a factor of 0.5 for normal faulting.

The model of *RC* includes the effects of the pore fluid pressure on the frictional resistance via the effective stress principle, the reduction in pore fluid pressure due to dilatancy, and the fluid flow to the slip zone in response to pressure reduction. The analysis suggests that dilatant hardening accompanying slip stabilizes the response (prevents an unbounded slip rate) if $A$ is not too large and the ambient (initial) pore pressure is sufficiently high. The magnitude of stabilization depends on the mode of faulting. For drained response (no pore pressure changes), slip is unstable if $A$ exceeds 1 for no normal stress changes, 0.65 for thrust geometry and 1.5 for normal faulting (The latter two cases assume a fault that makes a 30° angle with the maximum compressive stress and a friction coefficient of 0.6). When pore pressure effects are included, the critical value of $A$ depends on $\alpha$ (which is proportional to the fluid compressibility) and the dependence is strong for small $\alpha$. For $\alpha = 0.1$, the minimum value of $A$ for unstable

response is 8.1 for no normal stress changes, 7.5 for the thrust geometry, and 9.6 for normal faulting.

As discussed in *RC* pore pressure stabilization is limited by the ambient pore pressure: That is, the pressure reduction due to dilatancy cannot exceed the initial ambient pressure and the stabilizing effect diminishes with an increase of pore fluid compressibility as the pressure is reduced. The magnitude of the maximum pressure reduction is proportional to the product of the stress drop ($\tau_p$-$\tau_r$) and the friction coefficient ($m_0$). The magnitude of the maximum pressure reduction also depends on the faulting mode. It is greatest for thrust faulting and least for normal faulting. The ratio of the maximum pressure reduction for normal to thrust faulting varies from zero to 1 as $A$ varies from the minimum value needed to cause instability in drained response to infinity.

The dependence of stabilization on the ambient pore pressure suggests, as surmised by *Sibson* [1991] (also, see *Rice* [1992] and *Byerlee* [1990,1993]),that the response may be sensitive to the hydrologic conditions near the fault. Seals that cause a build-up of pore pressure would make possible greater stabilization. Conversely, a permeable structure would tend to prevent any stabilization. Furthermore, the hydrologic properties near the fault may be altered by slip, e.g., seals may be broken, causing a more complex response.

Although the *RC* model includes features likely to be of importance *in situ*, the model is rudimentary and makes a number of approximations in favor of simplicity. Among these are that the friction is rate (and state) independent and that the fluid flow to the slip surface can be approximated by the steady state relation. The predictions of a more elaborate treatment of these aspects would differ in detail, but likely not in nature, from those here. *RC* also approximate the pore fluid bulk modulus as constant down to some critical pressure near the liquid-vapor transition. They argue that this approximation is reasonable for liquid water and temperatures and pressures in the shallow crust. But, the presence of dissolved gases and elevated temperatures could make a more sophisticated treatment necessary. This is also suggested by the sensitive response with constant compressibility on the value of the compressibility when it is large.

*Acknowledgment.* I am grateful to the organizers for inviting me to the Workshop and to the U.S. Geological Survey and the National Science Foundation for financial support of work on which this paper is based.

## REFERENCES

Brace, W. F., B. W. Paulding and C. Scholz, Dilatancy in the fracture of crystalline rock, *J. Geophys. Res.*, **71**, 3939-3953, 1966.

Byerlee, J., Friction, overpressure and fault normal compression, *Geophys. Res. Lett.*, **17**, 2109-2112, 1990.

Byerlee, J., Model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology*, **21**, 303-306, 1993.

Martin, R. J., III, Pore pressure stabilization of failure in Westerly granite, *Geophys. Res. Lett.*, **7**, 404-406, 1980.

Rice, J. R., Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault, in *Fault mechanics and transport properties of rock: a Festschrift in Honor of W. F. Brace*, edited by B. Evans and T.-F. Wong, pp. 475-503, Academic Press, San Diego, Calif., 1992.

Rudnicki, J. W. and C.-H. Chen, Stabilization of rapid frictional slip on a weakening fault by dilatant hardening, *J. Geophys. Res., 93,* 4745-4757, 1988.

Sibson, R. H., Loading of faults to failure, *Bull. Seismol. Soc. Am.*, **81**, 2493-2497, 1991.

Teufel, L. W., Pore volume changes during frictional sliding of simulated faults, in *Mechanical Behavior of Crustal Rocks, Geophys. Monogr. Ser.*, vol. 24, edited by N. L. Carter, M. Friedman, J. M. Logan, and D. W. Stearns, pp. 135-145, AGU, Washington, D. C., 1981.

# Near-Tip Stress Fields for Dynamically Propagating Mode-II Fractures

Allan M. Rubin and Christian B. Parker
*Department of Geological and Geophysical Sciences, Princeton University, Princeton, NJ 08544*

Field observations of exhumed faults indicate that fault zones often have a complex internal structure consisting of nested zones of variously fractured and brecciated rock (Wallace and Morris, 1986; Chester et al., 1993). Deformation within these zones may exert a strong influence on the overall constitutive behavior of the fault. For example, Sleep and Blanpied (1992) hypothesize that porosity created during dynamic rupture lowers the fluid pressure within fault zones, and that subsequent (interseismic) pressure solution increases the pore pressure until the next earthquake is triggered. The observed deformation within fault zones may result from the process of fault formation, large-scale slip along non-planar faults, or individual earthquakes. The goal of this study is to help assess the deformation that might be due to dynamic rupture events.

The model adopted is an extension to Mode-II shear fractures of the Barenblatt cohesive zone model for tensile fractures (Palmer and Rice, 1973). As indicated in Figure 1, the shear stress at a particular point on the fault plane increases as the rupture front approaches, reaches a peak value $\tau_p$ at which slip begins, and then, after the rupture front passes, decreases with increasing slip to a residual fault strength $\tau_f$ after a critical slip distance is achieved. The model parameters are chosen such that the stress singularity due to the stress drop over the entire fault is cancelled by that due to the larger resisting stresses within the slip weakening zone. Several idealizations are made in order to make an analytic treatment possible: (1) The fault is planar. (2) The medium is homogeneous. (3) The rupture propagates at constant velocity. Points (1) and (2) are common to most studies of dynamic rupture propagation. Restriction (3) is not shared by boundary integral or finite difference techniques. The advantage of our approach lies in the fact that it enables one to examine the stresses off the fault plane analytically, at the same time that the tip propagation criterion is exactly met. Thus the question of at what point small-scale inelastic deformation may occur off the fault plane can be addressed in a sensible fashion. It has been suggested that such deformation limits the propagation velocity of natural ruptures to less than elastic wave speeds (Rice, 1980). A similar (quasi-static) model has been quite successful at explaining outcrop-scale deformation produced at the tips of propagating dikes (Rubin, 1993).

The necessary equations were derived from the plane strain solutions of Freund (1990) for a point load applied a fixed distance behind the tip of a steadily moving, semi-infinite crack. The length $l$ over which the residual fault strength $\tau_f$ is applied in the model is intended to represent the actual fault length in the Earth, but we restrict our attention to the stress field near the rupture front. It is expected that the constant-velocity solution will be a reasonable approximation for ruptures that move with a near-constant velocity for distances long enough that any elastic waves have moved well outside the region of interest. Among other things this implies that ruptures propagating near elastic wave speeds may have to travel distances considerably greater than the length of the slip-weakening zone in order for the constant-velocity calculations to be relevant. This is important because for constant propagation velocity the dynamic stresses for Mode-II ruptures do not increase dramatically except at velocities approaching the Rayleigh wave speed $C_R$.

In the plots to follow, the rupture front is located at the center of the plot and propagates to the right. Stress values are normalized by $(\tau_p-\tau_f)$. There is no change in normal stress or pore pressure along the fault plane due to linear poroelastic effects; thus $(\tau_p-\tau_f)$ is given by

$$(\tau_p-\tau_f) = (\mu_s-\mu_d)|\sigma_{yy}^o-p^o| + \mu_d \Delta p^{fault}$$

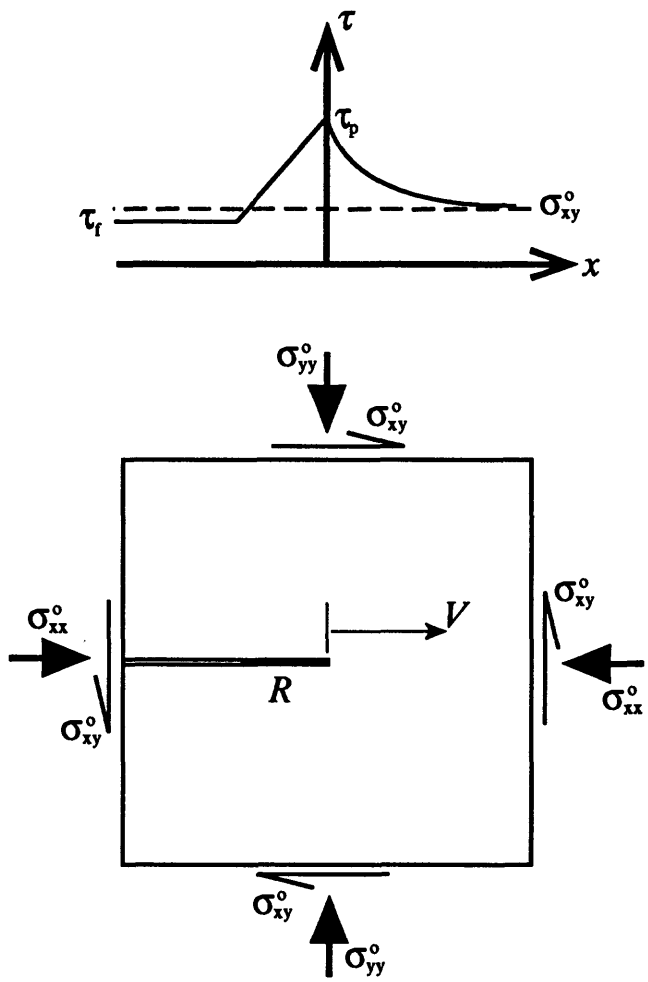



**Figure 1:** (bottom) Schematic diagram of model problem. (top) Shear stress $\tau$ along fault plane.

where the first term on the right is simply the difference between the static and dynamic coefficients of friction times the ambient effective normal stress, and the second allows for a decrease or increase in fault pore pressure $\Delta p^{fault}$ that might result from inelastic dilation or frictional heating. Compressive stresses and pore pressures are taken to be negative. Setting $\Delta p^{fault} = 0$, $\mu_s = 0.85$, and $\mu_d = 0.6$, for example, at 5 km depth an order of magnitude estimate of $(\tau_p - \tau_f)$ is 20 MPa for hydrostatic pore pressure and 10 MPa for an elevated pore pressure such that the "apparent" dynamic coefficient of friction (computed assuming hydrostatic pore pressure) is 0.2.

All distances are normalized by the length of the slip weakening zone $R$. In the limit that $R \ll l$, the resulting near-tip stress distribution is insensitive to $l$, and this is the case examined here. The stress intensity factor $K$ and relative slip $\delta$ a small distance $R$ behind the rupture front due to the uniform stress drop $(\sigma^o_{xy} - \tau_f)$ over $l$ are:

$$K = (\sigma^o_{xy} - \tau_f) \frac{2\sqrt{2}}{\pi} \sqrt{l}$$

$$\delta = (\sigma^o_{xy} - \tau_f) \frac{8}{\pi} \sqrt{l} \frac{1-\nu}{G} \sqrt{R} f_{II}$$

where $G$ is elastic shear modulus, $\nu$ is Poisson's ratio, and $f_{II}$ is a function of propagation velocity that is essentially 1 at moderate velocities but that increases without bound near the Rayleigh wave speed (Rice, 1980). Assuming for simplicity that within the slip weakening zone the fault strength decreases linearly with distance behind the rupture front, $K$ and $\delta$ due to the load acting within the slip weakening zone are

$$K = -(\tau_p - \tau_f) \frac{2}{3} \frac{2\sqrt{2}}{\pi} \sqrt{R}$$

$$\delta = -(\tau_p - \tau_f) \frac{1}{6} \frac{8}{\pi} \sqrt{R} \frac{1-\nu}{G} \sqrt{R} f_{II}$$

Setting the sum of the stress intensity factors to zero to eliminate the stress singularity yields

$$\frac{R}{l} = \frac{9}{4} \left( \frac{\sigma_{xy}{}^o - \tau_f}{\tau_p - \tau_f} \right)^2$$

$$R = \frac{1}{f_{II}} \frac{\pi}{4} \frac{1}{\tau_p - \tau_f} \frac{G}{1-\nu} \delta_c$$

where $\delta_c$ is the critical slip weakening distance. Thus $R \ll l$ provided $(\sigma^o_{xy} - \tau_f) \ll (\tau_p - \tau_f)$. For $(\tau_p - \tau_f) = 15$ MPa and $G/(1-\nu) = 30$ GPa, at zero velocity $R \sim 20$ m for $\delta_c = 1$ cm and 200 m for $\delta_c = 10$ cm. For a rupture velocity of $0.98C_R$, $f_{II} \sim 6.6$ and $R$ is reduced accordingly. Because a linear decrease in fault strength with distance behind the rupture front is not the same as a linear decrease with slip magnitude, the shear fracture energy is not $(1/2)(\tau_p - \tau_f)\delta_c$ but $(4/9)(\tau_p - \tau_f)\delta_c$ (Palmer and Rice, 1973), or about $10^5$ and $10^6$ J/m$^2$, respectively, for the two slip weakening distances above.

Figures 2a-2c show the change in near-tip shear stress $\Delta\sigma_{xy}$ for propagation velocities of 0.01, 0.90, and 0.98 times the Rayleigh wave speed. Slip is right-lateral, and compressive stresses and pore pressures are negative. The results do not vary greatly from the static case for rupture velocities less than $0.8C_R$. As pointed out by Freund (1979), $C_R$ represents the upper limit to physically plausible (constant) rupture velocities. Figures 2d-2f show $\Delta\sigma_{xx}$, $\Delta\sigma_{yy}$, and the change in pore pressure $\Delta p$ at a velocity of $0.98C_R$. The latter plot assumes that deformation is sufficiently rapid that undrained poroelastic response results. In this case $\Delta p$ is given by

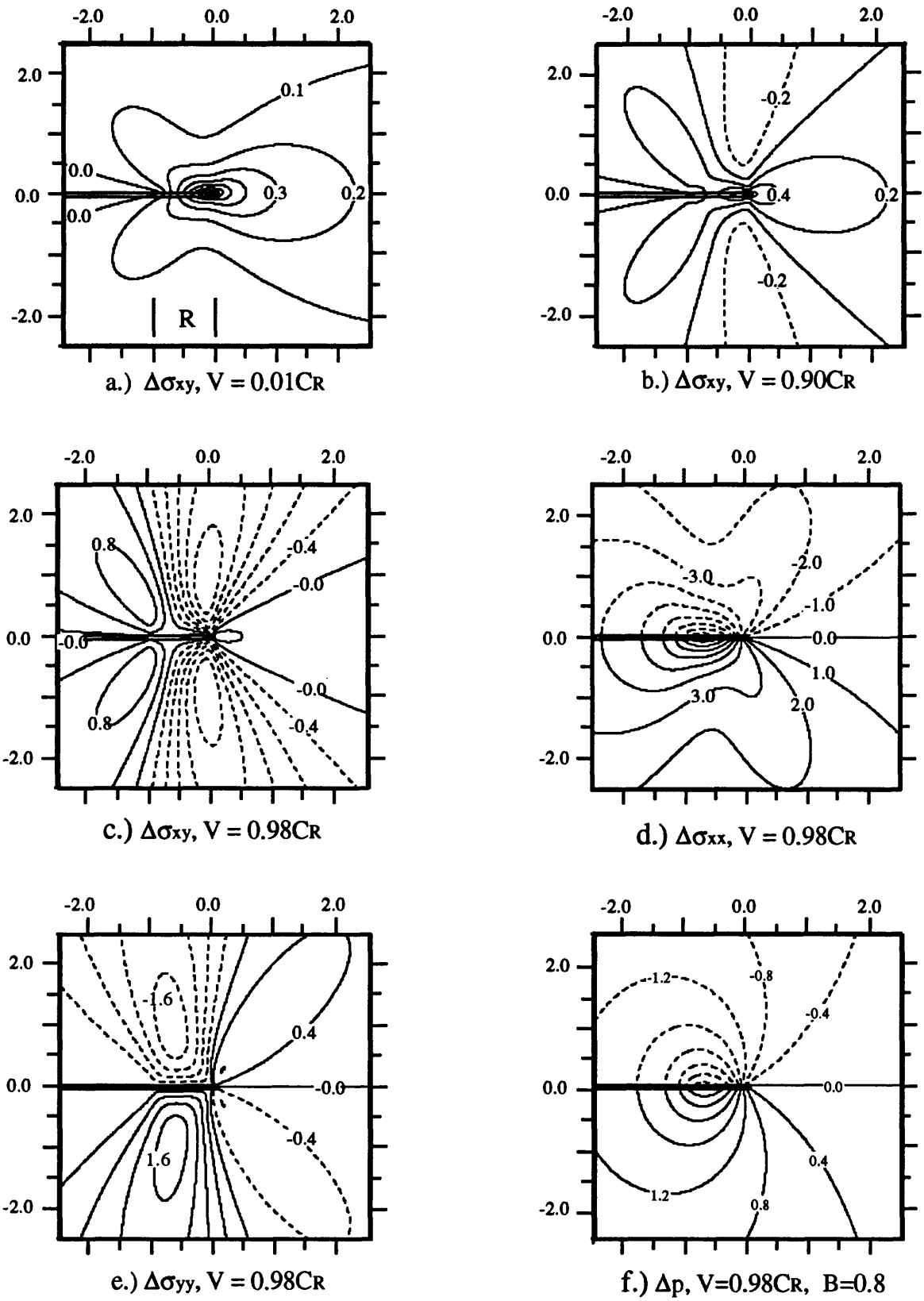



**Figure 2:** Contours of indicated stress values, normalized by ($\tau_p$-$\tau_f$).

$$\Delta p = B\frac{\Delta\sigma_{xx}+\Delta\sigma_{yy}+\Delta\sigma_{zz}}{3} = B\frac{(1+\nu)(\Delta\sigma_{xx}+\Delta\sigma_{yy})}{3}$$

where $B$ is Skempton's coefficient and the second equality results from plane strain deformation. Figure 2f assumes $B = 0.8$ and an undrained Poisson's ratio $\nu$ of 0.25. The computed step in pore pressure across the fault will become diffuse as the rupture front passes, as a pore pressure boundary layer develops with thickness of order $\sqrt{\omega t}$, where $\omega$ is hydraulic diffusivity. However, given typical values of $\omega$ for rock, this boundary layer thickness will be negligible for ruptures that propagate at near seismic velocities. Note that along the fault plane $\Delta\sigma_{yy} = 0$, and, with no inelastic dilation or frictional heating along the fault, from symmetry $\Delta p = 0$ as well.

The simplest failure criterion to evaluate is for slip on planes parallel to the fault, as this depends on the fewest number of independent parameters. Such failure may control the slip distribution within fault cores with sub-parallel walls. Slip is assumed to occur if

$$\sigma_{xy} > \mu_s|\sigma_{yy}-p|$$

Writing these stresses as the superposition of the ambient stress and the stress change due to rupture yields

$$\sigma^o_{xy} + \Delta\sigma_{xy} > \mu_s|(\sigma^o_{yy}+\Delta\sigma_{yy})-(p^o+\Delta p)| = \mu_s|(\sigma^o_{yy}-p^o)+\Delta\sigma_{yy}-\Delta p|$$

$$\sigma^o_{xy} + \Delta\sigma_{xy} > \mu_s|\sigma^o_{yy}-p^o| - \mu_s(\Delta\sigma_{yy}-\Delta p)$$

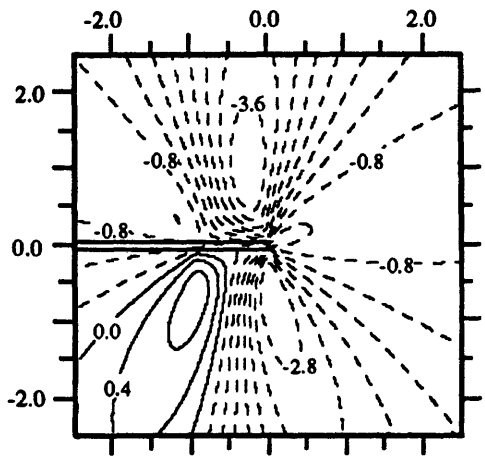
Because $\Delta\sigma_{yy}=\Delta p=0$ along the fault plane ahead of (and at) the rupture front, at the rupture front $\mu_s|\sigma^o_{yy}-p^o| = \mu_s|\sigma_{yy}-p| = \tau_p$, so failure occurs if

$$\sigma^o_{xy} + \Delta\sigma_{xy} > \tau_p - \mu_s(\Delta\sigma_{yy}-\Delta p)$$

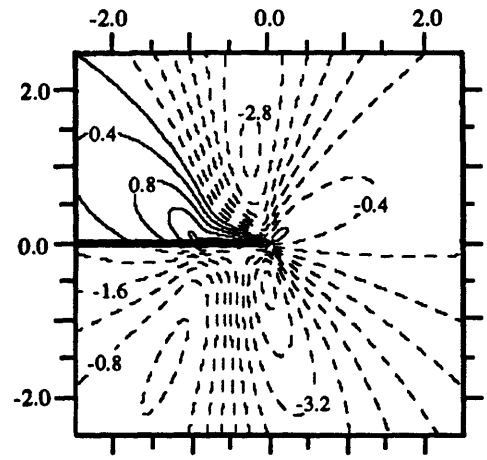
$$\Delta\sigma_{xy} > -\mu_s(\Delta\sigma_{yy}-\Delta p) + (\tau_p-\sigma^o_{xy})$$

The amount by which $\Delta\sigma_{xy}$ exceeds the right-hand side of the above equation is plotted in Figure 3a for dry rock, and in 3b for wet rock under undrained conditions, for $V=0.98C_R$, $B=0.8$, and $\nu=0.25$. Provided $R<<l$, $(\tau_p-\sigma^o_{xy})/(\tau_p-\tau_f) \sim 1$ and the normalized stress values are insensitive to the magnitudes of the ambient stresses. Positive values indicate that the slip criterion is exceeded locally, even while the same criterion is met at the fault tip. While the change in shear stress $\Delta\sigma_{xy}$ is symmetric across the fault (Fig. 2c), in the dry case slip is predicted on the side of the fault that moves away from the crack tip because $\Delta\sigma_{yy}$ is tensile (Fig. 2e). When the effect of pore fluid pressure is considered, however, this asymmetry is reversed, because the increase in pore pressure dominates the increase in $\Delta\sigma_{yy}$ (note the large compressive $\Delta\sigma_{xx}$ in 2d). The small region of positive values in front of the rupture in 3b suggests that the rupture might tend to migrate to the outside of the fault zone core, as long as potential failure planes are available for slip.

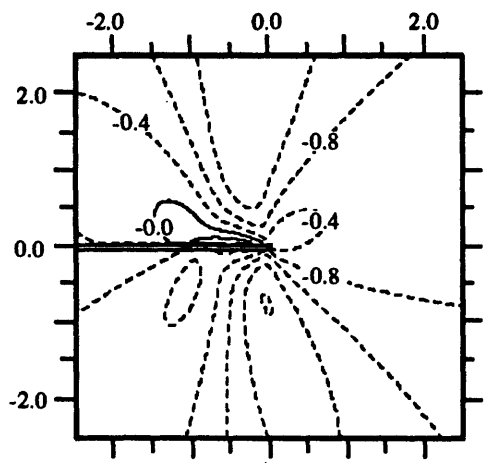
Figures 3c and 3d illustrate the amount by which the shear stress exceeds the static coefficient of friction times the effective normal stress on optimally oriented, rather than fault-parallel, planes. In this case, the normalized stress values depend upon the (normalized) magnitude of the stress parallel to the fault. In Figures 3c and 3f $\sigma^o_{xx}$ is chosen such that the fault zone is oriented optimally for slip at the dynamic coefficient of friction of 0.6; that is, the fault is about 30° to the direction of greatest compression S1. Failure is predicted over a modest region at $0.95C_R$ (Fig. 3c), and over a sizeable region at $0.98C_R$ (Fig. 3d). The similarity of the contours on the tensile (upper) side of the fault in 3d and 3b indicate that the optimal planes for slip are in fact close to fault-parallel; they are rotated about 5° to 15° counterclockwise from the fault over most of this



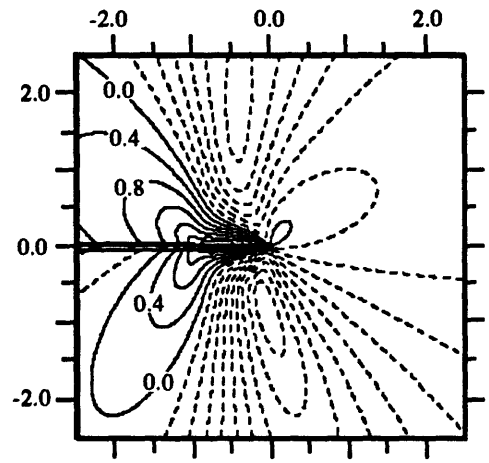

a.) Shear failure on fault-parallel planes, dry, V=0.98 $C_R$



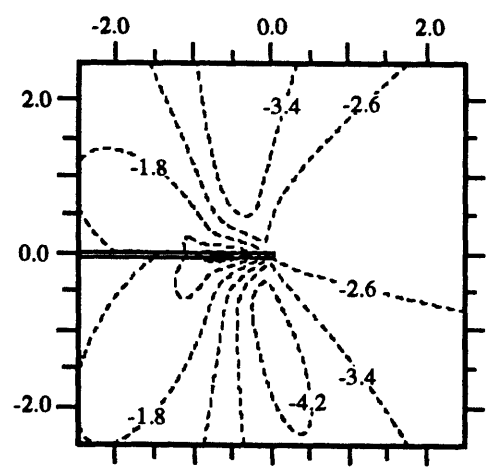

b.) Shear failure on fault-parallel planes, wet, V=0.98$C_R$, B = 0.8



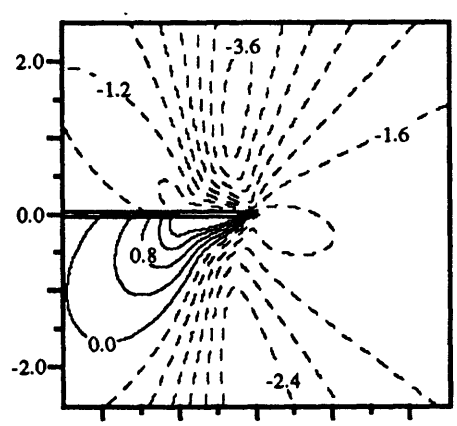

c.) Shear failure on optimal planes, V=0.95$C_R$, B=0.8




d.) Shear failure on optimal planes, V=0.98$C_R$, B=0.8




e.) Effective stress, V=0.98$C_R$, B=0.8, S1 30° to fault




f.) Effective stress, V=0.98$C_R$, B=0.8, S1 45° to fault

**Figure 3:** Contours of indicated failure criteria, normalized by ($\tau_p$ - $\tau_f$).

region. Within the zone of failure on the compressional side of the fault, the optimal planes are rotated by a similar amount clockwise.

In order to consider production of porosity or observed veins within fault zones (Chester et al., 1993), one must consider the relation between the pore fluid pressure and the least compressive stress. This is contoured in Figure 3e for the same parameters as in 3d. No excess pore pressure occurs for $V$=0.95$C_R$, and vein formation is predicted only in a very narrow region immediately adjacent to the extensional (lower) side of the slip weakening region for $V$=0.98$C_R$. This zone is about 0.07$R$ wide, with maximum excess pore pressures of about $0.7(\tau_p-\tau_f)$. Diffusion of pore pressure from the high pressure (upper) side of the fault could enhance fracturing over a limited portion of this thickness. The predicted vein orientation is about 70° from the fault. As pointed out by Rice (1980), as the Rayleigh wave velocity is approached $\Delta\sigma_{xx}$ increases without bound and tensile cracks would be predicted to be closer to perpendicular to the fault. For a lower $\sigma^o_{xx}$, that is, for a greatest compression direction closer to fault-perpendicular, veins could form at lower rupture propagation speeds. For $\sigma^o_{xx} = \sigma^o_{yy}$, so that the greatest compression is 45° to the fault, the magnitude and extent of the zone of excess pore pressure is about the same for $V$=0.95$C_R$ as it is in Figure 3e for $V$=0.98$C_R$. At $V$=0.98$C_R$ and $\sigma^o_{xx} = \sigma^o_{yy}$, the maximum excess pore pressure exceeds $3.4(\tau_p-\tau_f)$ immediately adjacent to the slip weakening region (Fig. 3f), where the stress orientation is such that tensile cracks would form at about 75° to the fault plane. While failure is predicted to occur over a region considerably larger than the slip weakening zone in 3f, tensile crack growth is self-limiting to the extent that rapid crack widening reduces the pore pressure.

Given that these elastic solutions neglect fault surface roughness and abrupt changes in rupture velocity that might also produce inelastic deformation, their usefulness hinges upon their ability to explain observed deformation within fault zones. Veins at high angles to faults may be diagnostic of failure by the mechanism explored here, and such features are worthy of closer examination. Other suggestions for future work include (1) comparison with the stresses near propagating Mode-III ruptures, and (2) comparison with numerical models such as that of Harris and Day (1993). This would serve the dual role of determining the range of applicability of the constant-velocity assumption used here, and assessing the accuracy of the numerical models.

## REFERENCES


Chester, F.M., J.P. Evans, and R.L. Biegel, Internal structure and weakening mechanisms of the San Andreas Fault, *J. Geophys. Res.*, *98*, 771-786, 1993.

Freund, L.B., The mechanics of dynamic shear crack propagation, *J. Geophys. Res.*, *84*, 2199-2209, 1979.

Freund, L.B., *Dynamic Fracture Mechanics*, 563 pp., Cambridge Univ. Press, Cambridge, 1990.

Harris, R.A., and S.M. Day, Dynamics of fault interaction: Parallel strike-slip faults, *J. Geophys. Res.*, *98*, 4461-4472, 1993.

Palmer, A.C., and J.R. Rice, The growth of slip surfaces in the progressive failure of over-consolidated clay, *Proc. R. Soc. Lond. A*, *332*, 527-548, 1973.

Rice, J.R., The mechanics of earthquake rupture, in *Physics of the Earth's Interior*, ed. E. Boschi, North-Holland, Amsterdam, p. 555-649, 1980.

Rice, J.R., Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault, in *Fault Mechanics and Transport Properties of Rocks*, eds. B. Evans and T.-F. Wong, Academic Press, London, p. 475-503, 1993.

Rubin, A.M., Tensile fracture of rock at high confining pressure: Implications for dike propagation, *J. Geophys. Res.*, in press, 1993.

Sleep, N.H., and M.L. Blanpied, Creep, compaction, and the weak rheology of major faults, *Nature*, *359*, 687-692.

Wallace, R.E., and H.T. Morris, 1986, Characteristics of faults and shear zones in deep mines, *Pure Appl. Geophys.*, *124*, 107-125, 1986.

# QUASI-STATIC SIMULATIONS OF EARTHQUAKES AND SLIP COMPLEXITY ALONG A 2D FAULT IN A 3D ELASTIC SOLID

Yehuda Ben-Zion and James R. Rice
Department of Earth and Planetary Sciences and Division of Applied Sciences
Harvard University, Cambridge, MA 02138

## ABSTRACT

Numerical simulations of slip instabilities on a vertical strike-slip fault in an elastic half-space are performed in a quasi-static fashion for various models belonging to two different categories. The first category consists of inherently discrete cellular fault models. Such are used to represent fault systems made of elements (modeled by numerical cells) that can fail independently of one another. Their quasi-independence is assumed to provide an approximate representation of strong fault heterogeneity, due to geometric or material property disorder, that can arrest ruptures at segment boundaries. The second category consists of models having a well defined continuum limit. These involve a fault governed by rate- and state-dependent friction, and are used to evaluate what types of property heterogeneity could lead to the quasi-independent behavior of neighboring fault segments assumed in the first category. The cases examined include models of a cellular fault subjected to various spatially complex distributions of static to kinetic strength drops, and models incorporating rate- and state-dependent friction subjected to various distributions of effective stress (normal stress minus pore pressure). The results indicate that gradual effective stress variations do not provide a sufficient mechanism for the generation of observed seismic response. Strong and abrupt fault heterogeneity, as envisioned in the inherently discrete class, is required for the generation of slip complexities, rough rupture fronts capable of radiating high frequency seismic waves, and a wide spectrum of event sizes. The large earthquakes in the discrete models occur on a quasi-periodic basis; however, the nature of the large events is highly variable. The results indicate that expectations for regular sequences of earthquakes and/or simple repetitive precursory slip patterns are unrealistic. The frequency-size (FS) statistics of the failure episodes simulated by the cellular fault models are approximately self-similar for small events, with $b \approx 1.2$ (the corresponding $b$-value of rupture area, $b_A$, is about 1). However, for events larger than a critical size the simulated statistics are, in general, strongly enhanced with respect to self-similarity. This is due to the fact that the stress concentrated at the edge of a rupture expanding in an elastic solid grows with the rupture size. Thus, when the fault properties (i.e., geometric irregularities) are characterized by a narrow range of size scales, the scaling of stress concentrations with the size of the failure zone creates a critical rupture area terminating the self-similar earthquake statistics. In such models, events reaching the critical size become (on the average) unstoppable and they continue to grow to a size limited by a characteristic model dimension. When, however, the system is characterized by a broad spectrum of size scales, the above phenomena are suppressed and the range of self-similar FS statistics is broad and characterized by average $b$- and $b_A$-values of about 1. The simulations indicate that local maxima in observed FS statistics correspond to dimensions of coherent brittle zones, such as the width of the seismogenic layer or the length of a fault segment bounded by barriers. Thus, observed FS statistics can be used to obtain information on crustal thickness and fault zone structure. The results suggest that the frequency of occurrence of moderate and strong earthquakes is enhanced with respect to self-similar distributions defined by small events, regardless of the shape of the observed statistics.

## INTRODUCTION

Fault zone properties governing seismic response include, among other things, the distributions of geometrical irregularities, material heterogeneities, temperature, and effective stress (normal stress minus pore pressure). Since the above distributions are complex functions of space (and time), the distributions of fault friction and earthquake stress drops can be rather involved.

One class of models for the evolution of stress and slip on faults employs constitutive stress-slip relations based on laboratory sliding experiments. These include models using slip-weakening constitutive law [e.g., Stuart and Mavko, 1979; Stuart et al., 1985; Stuart, 1986] and rate- and state-dependent friction [e.g., Dieterich, 1986; Tse and Rice, 1986; Horowitz and Ruina, 1989; Rice, 1993]. These models are attractive since they incorporate experimentally-derived descriptions of the frictional process that can be used to simulate stable sliding, catastrophic failure and, most importantly, the transition between these two modes of slip. However, it is not yet clear whether and how laboratory friction results, based on small rock samples with relatively smooth slip surfaces, correspond to properties of a natural fault system having a wide range of geometrical heterogeneities, as well as far larger slips (and slip rates) during instabilities and longer interseismic intervals for fault rehealing.

Another modeling approach, employing theoretical solutions from fracture mechanics, prescribes fault properties in terms of critical energy release rate, or critical stress intensity factor, characterizing cohesive material forces that resist sliding [e.g., Gao et al., 1991; Yamashita and Knopoff, 1992; Rice and Ben-Zion, 1993; Yamashita, 1993]. As discussed in Rice [1980], the description of fault properties in terms of critical energy release rate (or critical stress intensity factor) corresponds to a description based on laboratory-derived constitutive laws, in the limiting case where the "critical slip distance", characterizing the evolution of frictional properties with slip, is small compared to other model dimensions, and when the shear stress on the fault is relatively independent of slip rate or history for slips greater than the critical distance.

A third modeling approach, not addressing the physics of the transition in the frictional properties of sliding surfaces, simply prescribes various distributions of static/kinetic frictions or earthquake stress drops on a fault. This approach, formally corresponding to a case of zero critical slip distance and additional simplifications neglecting the dependence of fault strength on slip velocity, was used in the past [e.g., Andrews, 1980, 1981; Kagan and Knopoff, 1981; Kagan, 1982; Rundle, 1988; Frankel, 1991; Ben-Zion and Rice, 1993] and is adopted in much of the present work, in attempts to elucidate general characteristics of fault response to various complex distributions of properties. Ben-Zion and Rice [1993] argued that this third approach could provide a valid approximate treatment of strong geometrical or material property variations which can stop ruptures. Using cells of a numerical grid as an approximate representation of fault segments, delimited by such potentially arresting heterogeneities, the approach may be justified when the critical nucleation size $h^*$ of Rice [1993] is much smaller than fault segment size.

A key conceptual issue is whether the spatio-temporal complexity of fault slip can be attributed purely to fault dynamics, or whether nonuniformities in geometry, material properties, effective stress, and/or mode of loading are essential for understanding the observed complex fault slip. That is, are seismicity in the crust and turbulent flow in a fluid of uniform physical properties analogous as advocated, e.g., in Kagan [1992]? A number of recent theoretical studies indicate that the answer may be positive. Horowitz and Ruina [1989] generated complex spatial and temporal slip events in a homogeneous model as a direct consequence of rate- and state-dependent friction law, although they could do so only for carefully selected model parameters. Bak and Tang [1989], Ito and Matsuzaki [1990] and others modeled earthquake dynamics as a manifestation of self organized criticality in the crust, using cellular automata governed by a simple system evolution rule. Carlson and Langer [1989], Carlson et al. [1991], Brown et al. [1991], Christensen and Olami [1992] and others modeled fault dynamics with extended arrays of Burridge and Knopoff [1967] spring-connected blocks and pure velocity dependence of friction. However, with the exception of the "tuned" Horowitz and Ruina study, these (cellular automata and block-spring) models are problematic since they have system dimension the same as that of the fault,

rather than one greater, and they employ simplified stress transfer laws (i.e., nearest neighbor interactions) that do not scale with rupture size in a way suitable to represent faults embedded in elastic continua. Nevertheless, the Burridge-Knopoff models with suitable parameter choices do show complexity arising from Newtonian dynamics, possibly in a way unaffected by what Rice [1993] called the "inherent discreteness" (see below) of such models.

Rice [1993] and Ben-Zion and Rice [1993] used quasi-static numerical simulations to study mechanical instabilities along 2D fault zones embedded in 3D elastic solids. The work of Rice involved a smooth fault surface governed by rate- and state-dependent friction, while the work of Ben-Zion and Rice focused on cellular fault structure governed by static/kinetic friction. Rice [1993] found that there are fundamental differences between models of fault instabilities using separate classes of constitutive laws. When the constitutive stress-slip relation includes a finite critical slip distance over which the strength of failing elements evolves continuously, i.e., a slip weakening or state transition slip distance, the model has a well defined continuum limit which can be approached for small enough numerical cells. When, however, the strength of failing elements drops discontinuously with slip (as in the cellular automata models mentioned above and the work of Ben-Zion and Rice), or is dependent only on the velocity of slip such that no finite distance is included in the constitutive law (as in the cited block-spring simulations) the model is inherently discrete. In models with a continuum limit, suitably refined cells can fail only as members of a cooperating group, whereas in inherently discrete models cells can fail individually regardless of how small they are. When a continuum limit exists, the cell size $h$ of a suitably refined grid must be small compared to a nucleation size of slip instabilities $h^*$ (see the second analysis section for additional discussion and definition of $h^*$). The results of Rice [1993] indicate that the generic response of models with a continuum limit, using suitably refined cells and gradual variations of material properties, is a simple limit cycle of repeated earthquakes without a Gutenberg-Richter (GR) spectrum of small events. In contrast, the generic response of inherently discrete systems gives a spectrum of event sizes with some range of self-similar scaling. We note, however, that the work done so far does not rule out the possibility that proper elastodynamic solutions of smooth fault models with a continuum limit would lead to residual stress heterogeneity that can sustain complex seismic response. This possibility is discussed by Madariaga and Cochard [1992], although not for a model which, at least when analyzed quasi-statically, has a continuum limit.

In this work we present results based on two modeling approaches. In the first part of the analysis we employ the simulation procedure of Ben-Zion and Rice to study the distribution of slip, the shape of rupture fronts, and the frequency-size (FS) statistics of earthquakes in a discrete cellular model of the central San Andreas fault (SAF). The use of a discrete cellular model is motivated by numerous observations which indicate that crustal rocks are massively fractured and fault zone geometry is highly irregular. Examples include inter-event distance statistics of earthquakes [Kagan, 1991], profilometry measurements along exhumed fault surfaces [Power et al., 1988; Power and Tullis, 1991], characterization of mapped SAF traces [Aviles et al., 1987; Okubo and Aki, 1987], and analysis of geophysical logs from the 3.5 km deep Cajon Pass borehole [Leary, 1991]. Results from these and other studies show that crustal fractures and earthquakes are characterized by heterogeneities that cover a wide range of size scales. In the presence of fault bends, forks and step-overs, complex fault walls topography, and/or complex pore pressure variations, a fault segment may fail independently of an adjoining region in a way prohibited by the large stress concentration at the front of a slip event spreading along a smooth surface. We thus interpret the cells of our inherently discrete fault model as representing quasi-independent fault segments of size (equal cell size $h$, here of the order of 0.5 km) that is large compared to $h^*$. Additional motivation for using a discrete cellular fault structure to model seismic events can be found in Ben-Zion and Rice [1993].

The cases examined in the first section of the analysis cover various degrees of fault heterogeneity. In the least heterogeneous case, the computational grid has uniform properties and model heterogeneity comes only from the inherently discrete cellular fault structure and the assumed loadings. In more heterogeneous cases, the computational grid contains property variations. The results show spatio-temporal complexity in all cases; however, strong model fault heterogeneity is needed for the generation of displacement profiles compatible with geodetic and

seismological observations, and rough rupture fronts capable of radiating high frequency seismic waves. As noted earlier, strong fault heterogeneity is unavoidable in the presence of geometrical irregularities and substantial variations in the lithology and pore pressure of fault zone rocks. (Rice [1992a, 1993] found that 10% frictional property variations are ineffective in stopping rupture, although severe changes by factors of 5 to 10 can do so as can, of course, a transition to inherently stable velocity strengthening; see also the second analysis section of the present work.)

The FS statistics of the failure episodes simulated along our 2D cellular faults are self-similar for small events with $b \approx 1.2$ and $b_A \approx 1$, where $b$ and $b_A$ are $b$-values based on magnitude and rupture area, respectively. For events larger than a critical size, however, the simulated statistics are strongly enhanced with respect to self-similar distributions defined by the small events. This is due to the fact that the stress concentrated in a 3D elastic solid at the edge of an expanding rupture grows with the rupture size. Thus, beyond a critical rupture area events tend to grow to a size limited by a characteristic model dimension. This phenomenon is suppressed for a cellular fault model consisting of weak segments (e.g., overpressurized fault regions) of variable size, separated by strong boundaries (e.g., permeable zones with hydrostatic pressure). For other cellular fault models, the supercritical rupture growth results in FS statistics showing local maxima for events having rupture areas corresponding to the upper layer thickness and/or the entire computational grid. Similar local maxima are present in FS statistics of earthquakes observed along given faults [Singh et al., 1983; Schwartz and Coppersmith, 1984; Davison and Scholz, 1985; Wesnousky, 1993] (but see also Kagan [1993]) and in seismogenic zones with approximately uniform properties [Main and Burton, 1984; Main, 1987; Main and Burton, 1989; Trifu and Radulian, 1991].

In the second part of the analysis we use the simulation procedure of Tse and Rice [1986] and Rice [1993] to study the distribution of slip along a fault governed by rate- and state-dependent friction, and subjected to various assumed distributions of pore pressure. Our aim is to use the more rigorous constitutive framework to learn what types of property heterogeneity on a smooth fault can induce responses similar to observations and those generated by our discrete cellular models. As in Rice [1993], all cases calculated with cell size small enough to properly represent the underlying continuum model lead to periodic large earthquakes. A very strong pore pressure heterogeneity is needed to simulate in these models even a modest amount of complexity. This suggests that pore pressure variations in fault zone rocks are not the main mechanism responsible for the observed complex patterns of seismic slip, unless processes such as those proposed by Byerlee [1993] and Sleep and Blanpied [1992] create fault zone compartments with extreme pressure variations capable of stopping seismic ruptures of various size. Also in agreement with Rice [1993], calculations done with strongly oversized cells, so that the model is effectively discrete, show features resembling the response of the discrete cellular model examined in the first part of the analysis and observed seismic behavior. The results of the second analysis section support the hypothesis that discrete fault models, accounting for 3D continuum elasticity, may provide a valid approximate representation of fault zones having geometric disorder characterized by size scales much larger than the nucleation size of slip instabilities.

The simulations of Rice [1993], Ben-Zion and Rice [1993] and the present work show irregular sequence of earthquakes and a great diversity in the failure mechanism of model events. The results indicate that strategies for short-term earthquake prediction can not be based on simple repetitive precursory patterns such as accelerated microearthquake slip, although such occurs before some of the large model events. The studies suggest that the spatio-temporal complexity of fault slip is related to strong fault zone heterogeneities, probably due to geometric irregularities, capable of stopping ruptures of various sizes. We note that our results are based on quasi-static simulations with only approximate inclusion (e.g., by radiation damping term or allowance for dynamic overshoot) of elastodynamic effects.

## ANALYSIS

### *Modeling Based on Discrete Cellular Fault Zone Structure*

Figure 1 shows a model of the central SAF from Ben-Zion and Rice [1993]. The model consists of an elastic half-space with a vertical half-plane fault; the half-space has a 17.5 km thick brittle upper crust over a lower crust and upper mantle region where stable sliding occurs. The fault region in the upper crust contains large earthquake patches (regions I and IV) representing the rupture zones of the 1857 and 1906 M8 events. On these patches we impose a staircase slip history with a recurrence time of 150 yr. Everywhere along the lower crust and upper mantle (region V) and in the upper crust fault segment that is distant from Gold Hill (GH) by more than 70 km (region III) we impose slip at a constant plate rate $V_{pl}$ = 35 mm/yr. The upper crust fault segment $0 \le x \le 70$ km, $0 \le z \le 17.5$ km comprises a computational grid where space and time evolution of stress and displacement fields due to the imposed slip are calculated using a variant of static/dynamic friction law and 3D continuum elasticity. Different distributions of stress drops on failing computational cells are used to model fault regions with variable frictional properties. One model configuration, shown explicitly in Figure 1, is that of a fault containing a large rectangular region having different properties than the rest of the fault and designated as the "Parkfield asperity".

Consider a failure process involving a fault segment represented by a single cell of the computational grid. When the stress $\tau$ on the fault segment (cell) reaches a static strength $\tau_s$, failure occurs and the segment slips at a dynamic stress level $\tau = \tau_d$ ($<\tau_s$) until the rupture is blocked by the segment borders. Then, slip on the cell is brought to a halt, leaving on the failed segment an arrest stress $\tau_a$ ($<\tau_d$). If the stress transferred from the failed segment increases the stress at other fault regions to their failure thresholds those regions fail, causing additional stress transfers which, in turn, may induce or reinduce more slip events. We assume that reinitiation of slip on an already failed cell occurs when $\tau \ge \tau_d$ there. Thus $\tau_s$ is the failure strength of a segment which has not yet slipped in a composite event, $\tau_d$ is the failure strength of a segment which has slipped in an earlier sub-event, and $\tau_a$ is the stress remaining on a cell just after it has slipped, and before stress transfers from other segments failing at the same time have occurred. The static strength, dynamic strength, and arrest stress are related to each other as $(\tau_s - \tau_a)/(\tau_s - \tau_d) = D$, where $D$, set in this work to 1.25, is a dynamical overshoot coefficient. The difference $\tau(i, j) - \tau_a(i, j)$ gives a transient stress drop on cell $(i, j)$ during a composite failure episode, where $i$ and $j$ are, respectively, cell indexes along the horizontal and vertical directions and $\tau(i, j)$ is the stress just before failure. The final stress drop on cell $(i, j)$ is usually less than $\tau(i, j) - \tau_a(i, j)$ due to subsequent failures of other cells, but is always greater or equal to $\tau(i, j) - \tau_d(i, j)$.

The evolution of stress and displacement fields in the computational region is calculated using the quasi-static procedure outlined in Ben-Zion and Rice [1993]. The stress transfer due to incremental tectonic loadings and failing grid cells is computed by a discretized form of a boundary integral equation, based on the static solution of Chinnery [1963] for dislocations in a 3D elastic half-space. The results are independent of the absolute values of $\tau_s$, $\tau_d$, and $\tau_a$, since only the stress drops enter model calculations. We model differences in fault properties by assigning different levels of $\tau_s - \tau_a$ to different fault regions. The simulation procedure generates large ruptures as aggregates of small sub-events on quasi-independent fault segments, on each of which there are initiation, propagation and arrest events, and possible reinitiations or repeated reinitiations. The most important model parameters are the number and size of grid cells, the distributions of $\tau_s - \tau_a$, and the recurrence interval and failure timing of the 1857 and 1906 patches.

The calculations are done using $128 \times 32 = 4096$ square cells having a dimension of about 550 m. This is close to the size of the slip patches (dimension of about 200 m) that generate repeated microearthquakes at Parkfield [Antolik et al., 1991]. We begin all simulations with a 150 yr of model-conditioning analysis during which the 1857 and 1906 earthquake patches (regions I and IV in Figure 1) are locked, and the stably-sliding zones (regions III and V in Figure 1) move by the amount $V_{pl}\, t$ = 35 mm/yr · 150 yr = 5.25 m. The fault configuration at the end of the model-conditioning period provides nonzero heterogeneous initial states of stress and displacement

for the continuing analysis. At $t = 0^+$ yr into the analyzed earthquake cycle we impose 5.25 m of right-lateral slip (equal to the accumulated plate motion in 150 yr) on the 1857 and 1906 earthquake patches. For simplicity, both earthquake patches are moved simultaneously. We note that in the present modeling the computational grid is largely "shielded" from the 1906 earthquake patch by the presence of the uniformly-sliding region III.

We consider five cases of prescribed property distributions. (1) A fault with a large rectangular model Parkfield asperity. We assume that the static strength is $\tau_s = 40$ bar (4 MPa). In the region $0 \le x \le 25$ km, 5 km $\le z \le 10$ km (Parkfield asperity), $\tau_a = 0.1\tau_s$. In the other fault regions, $\tau_a = 0.9\tau_s$. The spatial extent of the model asperity is compatible with geodetic data [e.g., Harris and Segall, 1987] and microearthquake distribution [e.g., Malin et al., 1989] along the central SAF. The assumption of different stress drops in the asperity and nonasperity regions is compatible with the observations of O'Neill [1984] and Malin et al. [1989]. (2) A second example of an explicit model Parkfield asperity (Figure 2a). In the region $0 \le x \le 25$ km, 5 km $\le z \le 10$ km, $\tau_a = 0.1\tau_s \pm 0.1\tau_s$ uncorrelated random fluctuations. In the other fault regions, $\tau_a = 0.8\tau_s \pm 0.1\tau_s$ uncorrelated random fluctuations. Here the static strength is $\tau_s = 60$ bar. (3) A homogeneous fault with $\tau_s = 60$ bar and $\tau_a = 0.8\tau_s$. (4) A fault with a fractal property distribution (Figure 2b). The static stress is $\tau_s = 60$ bar. The arrest stress is $\tau_a = f\tau_s$, where $f$ is taken from a fractal distribution generated by a program outlined in Brown [1993]. The distribution is characterized by a fractal dimension 2.3, mean value 0.5, and standard deviation 0.2. (5) A fault consisting of "weak" (low stress drop) segments of variable size, separated by "strong" (high stress drop) boundaries (Figure 2c). Cells along the free surface are selected randomly with a probability of 1/5. The selected cells initiate one-cell-thick segment boundaries; these are extended to depth by 2D random walks. In cells belonging to segment boundaries, $\tau_a = 0.1\tau_s$. In other cells, $\tau_a = 0.9\tau_s$. The static strength is $\tau_s = 100$ bar. This case is taken as an example of the hydrological fault zone models of Byerlee [1990; 1993] and Rice [1992b]; the weak segments represent overpressurized regions, possibly delimited by very low permeability seals, while the strong segments represent more permeable zones with hydrostatic pore pressure.

Figures 3a and 3b show time evolution of slip along the SAF at a depth of 8.75 km for cases 2 and 5 above. In Figure 3a and in calculations corresponding to case 1 (not shown), where the assumed properties account explicitly for a large "Parkfield" asperity, slip in the region $0 \le x \le 25$ km is accommodated in the form of large steps separated by periods of little or no activity, while slip in the region $x > 25$ km is accommodated in a nearly continuous fashion. These results are compatible with geodetic data and earthquake distribution along the central SAF. On the other hand, in Figure 3b and in calculations corresponding to the property distributions 3 and 4 (not shown), time evolutions of slip in the region $0 \le x \le 25$ km are not clearly distinguishable from slip histories at other fault segments.

Figure 4a shows a typical example of a rupture area during a simulation corresponding to case 3 above. The results demonstrate that when the fault properties are smooth functions of space, the shape of the rupture front is also smooth. When, however, the fault properties vary rapidly in space, as in cases 1, 2, 4, and 5, the stress field along the rupture front is characterized by large spatial variations and a typical shape of the rupture front is rough. This is illustrated in Figure 4b for the property distribution of case 5. The results of Figures 3 and 4 indicate that strong fault heterogeneity is required, at least in this type of quasi-static models, for the generation of displacement fields compatible with geodetic and seismological observations, and rough rupture fronts capable of radiating high frequency seismic waves. There may also exist elastodynamic reasons for roughness of a rupture front.

Figure 5 shows cumulative rupture area as a function of time along the asperity region $0 \le x \le 25$ km, 5 km $\le z \le 10$ km; the property distribution is given by case 1 above. We note that the character of the large events varies from one earthquake to the other. The event at about 180 yr is preceded by a period of gradual rupture formation; the events at about 190, 220, and 240 yr are abrupt; the events at about 210 and 230 yr resembled "slow" earthquakes; the events at about 200 and 260 yr are preceded by "foreshocks"; and the event at about 280 yr is followed by an "aftershock". The diversity in the failure mechanism of the model asperity may be a manifestation of the many paths toward an equilibrium (low energy) state, available to a complex many-degrees-

of-freedom system. The results of this modeling indicate that expectations for simple and regularly occurring diagnostic precursory patterns, such as accelerated microearthquake slip, are unrealistic; such patterns may occur before some failures but they are not very indicative.

Figure 6a shows cumulative frequency-size statistics of rupture area and earthquake magnitude in the computational grid $0 \le x \le 70$ km, $0 \le z \le 17.5$ km during 150 analysis yr for the stress drop distribution of case 1. The units of rupture area are km$^2$; earthquake magnitudes are calculated using the moment-magnitude relation of Hanks and Kanamori [1979]; log = $\log_{10}$. There are three clear event groups in the figure: small events characterized by $b \approx 1.2$ and $b_A \approx 1$ ($b$ and $b_A$ are $b$-values based on earthquake magnitude and rupture area, respectively), intermediate size events having $b$, $b_A < 0.5$, and a final steep fall off. A region $b$, $b_A \approx 1$ is commonly observed and simulated, indicating that small events are self-similar. The intermediate region results from a supercritical rupture growth beyond a critical event size. This is better seen in Figures 6b where we show corresponding noncumulative FS statistics of rupture area. The small events fall more-or-less on a self-similar curve, but beyond a rupture area of about 80 km$^2$ events usually continue to grow to a size that is limited either by the upper layer thickness ($\approx$ 17.5 x 17.5 km$^2$) or the entire model dimensions ($\approx$ 70 x 17.5 km$^2$). As discussed in Ben-Zion and Rice [1993], the deviation of the simulated FS statistics from a power law (self-similar) behavior is a direct consequence of continuum elasticity. More specifically, it is due to the fact that the stress concentrated in an elastic solid at the edge of an expanding rupture grows with the size of the failure zone. Thus, when the assumed fault properties are characterized by a narrow range of size scales (representing a narrow range of geometric disorder), the scaling of stress concentrations with the rupture size introduces a critical event size terminating the self-similar earthquake statistics. In such systems events reaching the critical size become (on the average) unstoppable and they continue to grow to a size limited by a characteristic system dimension. We note that the simulated FS statistics are compatible with the observed Parkfield seismicity, where the rate of occurrence of five M6 Parkfield earthquakes in a century is strongly enhanced with respect to a self similar line defined by the small events.

The statistics of Figure 6 are generated by a model with a simple property distribution and two characteristic size scales, namely the size of the computational cell and the size of the model asperity. Figures 7a and 7b show, respectively, cumulative and noncumulative FS statistics for the fractal property distribution of case 4 (Figure 2b). Since the property distribution, although more complex than the model leading to Figure 6, is characterized by a single size scale, the range of self-similar statistics in Figure 7 is terminated at a critical rupture area similar to that of Figure 6. However, the highly heterogeneous (fractal) property distribution in the model corresponding to Figure 7 simulates barriers that can stop events in various stages between the critical rupture area and the size of the computational grid. Thus, the statistics of Figure 7 do not show the clear local maximum observed in Figure 6 for events with rupture area the size of the upper layer thickness. Instead, Figure 7 exhibits a general enhancement of seismicity beyond the critical rupture area and a single maximum for events the size of the whole system. Figures 8a and 8b show FS statistics for the property distribution of case 5, where the model represents weak fault segments of variable size and strong intervening barriers (Figure 2c). Here, the assumed fault properties are characterized by many size scales and, consequently, the range of self-similar FS statistics is broader than those of Figures 6 and 7. The average $b$- and $b_A$-values of the FS statistics in Figure 8 are about 1. We note that owing to the scaling of stress concentrations in elastic solid mentioned above, the rate of occurrence of intermediate and large size events in all of Figures 6-8 is enhanced with respect to self-similar lines defined by small events, regardless of the shape of the simulated statistics.

*Modeling Based on Rate- and State-Dependent Friction with Nonzero Critical Slip Distance*

We now describe some calculations in the framework of Tse and Rice [1986] and Rice [1993], with the aim of understanding how pore pressure distributions affect the earthquake cycle. Particularly, we examine how severe heterogeneities of pore pressure, $p$, may provide barriers adequate for stopping some ruptures, and hence lead to a complex seismic response. The results are also used to examine possible connections between our discrete and continuum models. Such

connections are important since the numerical difficulties of accurately implementing the rate- and state- dependent friction are formidable, and thus it is advantageous to use simpler compatible modeling approaches when appropriate. We limit attention here to distributions which are constant in time, and hence do not evaluate such mechanisms as proposed by Sibson [1992] and Sleep and Blanpied [1992] where $p$ is time dependent. All the reported results are based on analysis of the model shown in Figure 9, from Rice [1993], in which slip is imposed at $V_{pl}$ = 35 mm/yr on the portion of a half-plane fault lying below 24 km in an otherwise homogeneous elastic half-space. Slip $\delta$ is calculated on the fault section between 0 and 24 km depth, and over that range a Dieterich-Ruina type of rate-and state-dependent constitutive law applies. Thus, over the depth range 0-24 km the shear stress $\tau$ satisfies

$$\tau / (\sigma_n - p) = f_0 + a \ln [(d\delta / dt) / V_{pl}] + b \ \Psi$$
$$d\Psi / dt = - [(d\delta / dt) / L] \ \{\Psi + \ln [(d\delta / dt) / V_{pl}] \}. \qquad (1)$$

Here $\sigma_n$ is the normal stress (equated to depth times 28 MPa/km), $f_0$ is a constant which is irrelevant for the results to be shown, $L$ is the critical slip distance for state evolution, and $a$ and $b$ are functions of temperature which are converted to functions of depth just as in Rice [1993, Figure 3], based on the data of Blanpied et al. [1991] for granite under hydrothermal conditions. The assumed fault properties correspond to velocity strengthening regimes ($a > b$) for the depth sections 0-3 km and below 14 km, and a velocity weakening regime ($a < b$) elsewhere; in all calculated cases $b - a = 0.004$ over depths between about 4 and 13 km. The fault zone in the depth range 0-24 km is divided into numerical cells (Figure 9), and the system of equations is completed by writing from elastic theory

$$\tau_i = - \sum_j k_{ij} (\delta_j - V_{pl} \, t) - (\mu / 2c) \, d\delta_i / dt. \qquad (2)$$

The subscripts in (2) refer to cell indices and $k_{ij}$ is calculated from the Chinnery [1963] static solution as in the previous section. The last factor on the right side of (2) is a seismic radiation term with a damping coefficient $\mu/2c$ ($\mu$ = shear modulus = 30 GPa, $c$ = shear wave speed = 3 km/sec). As discussed by Rice [1993], this term provides an approximate representation of inertial effects in an elastic continuum and it allows solutions to exist during instabilities, when purely quasi-static calculations lead to unbounded slip velocity.

While Figure 9 suggests slip variation with depth and distance along strike, $\delta = \delta(x, z, t)$, we simplify matters here by constraining the slip to vary only with depth, i.e., $\delta = \delta(z, t)$. We discuss results of six calculations denoted as cases I to VI. Cases I to V are done with a cell height $h$ that is small enough to assure that the calculations represent the continuum limit of the rate- and state-dependent friction law (1). Case VI illustrates results for an oversized $h$, i.e., a discrete system; as discussed in the introduction, the results for this case are very different in character from the results of cases I to V representing continuum systems. Rice [1993] defined a nucleation size $h^*$ as the cell size $h$ such that the "single-cell stiffness" $2\mu/\pi h$ becomes equal to the "critical spring stiffness" $(b - a) (\sigma_n - p) / L$; the equality marks a border between stable and unstable sliding for velocity weakening behavior of the adopted friction law.

In cases I to V we use 512 cells through the 24 km thickness, so that $h$ = 0.047 km. The critical slip distance $L$ is variable from cell to cell, but it is never smaller than 1 mm. In velocity weakening regions $L$ is chosen larger as necessary to make the local value of $h^* = 2L \, \mu/\pi \, (b - a) (\sigma_n - p)$ coincide with $8h$ = 0.375 km. More specifically, in cases I to V, $L$/mm = max [1, 19.6 ($b - a$) ($\sigma_n - p$) / MPa]. In contrast, for case VI we take 64 cells through the thickness, so that $h$ = 0.375 km, and we choose $L$ so that in velocity weakening regions $h^* = h/4$, unless such requires an $L$ smaller than 1 mm (this is again the minimum allowed value). Since the nucleation size in case VI is much smaller than cell size, the corresponding model behaves as an inherently discrete system with response features comparable to those of our cellular model having the simplified static/kinetic friction law and $L = 0$.

Figure 10 shows various pore pressure distributions used in the following calculations. The thin line in each of the four panels corresponds to a hydrostatic pore pressure (= depth times 10 MPa/km). The first example, designated as case I, simply assumes $p$ = hydrostatic. This case, like all other calculated examples except case VI, leads to a periodic slip history. This is shown in Figure 11a where we plot slip versus depth; slip curves are drawn every 5 years and the total 14000 mm range of the slip axis corresponds to 400 years at 35 mm/yr. The recurrence interval for this case is about 140 yr.

Case II involves the pore pressure distribution of Figure 10a; this consists of a hydrostatic pressure gradient above and below 8 km, a depth where a severe pressure step elevates $p$ to 10 MPa below $\sigma_n$. This case represents a narrow but very low permeability layer that caps a deep high-$p$ zone below 8 km. The resulting slip history is shown in Figure 11b and, quite remarkably, even this very strong heterogeneity in $p$ is insufficient to stop ruptures and generate a complex sequence of events. Slips are again shown every 5 years, but now over a shorter total slip range of 6125 mm corresponding to 175 yr of model evolution at 35 mm/yr. Since the pore pressure in this case is generally higher than in the previous example, the recurrence interval is shortened to 65 yr.

Case III involves the pore pressure distribution of Figure 10b. Here the gradient begins as hydrostatic but changes to lithostatic [$dp/d(\text{depth}) = d\sigma_n/d(\text{depth})$] such that $p$ is offset by 50 MPa from $\sigma_n$ beyond some shallow depth range. This type of distribution was considered by Rice [1993] as an example for a pore pressure with asymptotically lithostatic gradient, shown by Rice [1992b] to result in a fault zone that is over-pressurized at depth and has a permeability which decreases strongly with increasing $\sigma_n - p$. The slip history is shown in Figure 11c, where now there are 1 yr intervals between the curves and the 3500 mm slip range corresponds to 100 yr. Here the recurrence interval is further shortened, in accord with the generally higher $p$, to 41 yr.

Case IV corresponds to the $p$ distribution of Figure 10c; after beginning with a hydrostatic gradient, $p$ jumps at 6 km (e.g., a horizon of very low permeability), then continues between 6 and 10 km with a lithostatic gradient such that $p$ is offset by only 10 MPa from $\sigma_n$, and finally continues with a hydrostatic gradient below 10 km. This pressure distribution is finally sufficient to induce some modest complexity in the earthquake history, as shown in Figure 11d; curves are again drawn every year and the slip range corresponds to 100 yr. We now see that moderate events, rupturing below 6 km and inducing nearly coincident slip above 4 km, occur between larger events that rupture the entire brittle crust. The recurrence interval for the large events is 38 yr and the moderate events occur 16 yr into the large event cycle.

Case V corresponds to the four-step $p$ distribution of Figure 10d. This might be thought of as a discontinuous version of the distribution in Figure 10b, in which permeable fault compartments, each 4 km deep and with hydrostatic gradient, are separated by very low permeability horizons at 4, 8, 12 and 16 km; each step raises $p$ to 10 MPa below $\sigma_n$. The slip history is shown in Figure 11e, once again with 1 yr intervals and slip range equivalent to 100 yr. Now a complex sequence of earthquakes occurs. The large events have a 52 yr recurrence interval, but around 1 yr into their cycle the compartment between 12 and 16 km ruptures as a large aftershock. Then, 17 yr into their cycle, a complex rupture occurs causing slip, or rapid afterslip, in all 4 compartments, but with nodes at the bottom of the most shallow compartment and, surprisingly, 3/4 of the way down in the compartment below. At 30 yr, the compartment between 12 and 16 km ruptures again. This case shows how strong heterogeneities of $p$, distributed over a fault zone, can induce complexities of rupture, although the matter is not very simple since the much stronger pressure step in Figure 10a (Case II) was insufficient to induce complexity. One viable (but still unproved) working hypothesis is that the expedient of using oversized cells, or setting $L = 0$ like in the static/kinetic analysis discussed earlier, may represent the type of heterogeneity effects modelled in case V, and other heterogeneities (presently too difficult to analyze) involving strong geometric disorder in the fault zone.

Figure 12 shows results for Case VI, describing a fault model with strongly oversized cells. The calculations are based on the $p$ distribution of Figure 10b, but general characteristics of the response with oversized cells do not seem strongly sensitive to details of the $p$ distribution (e.g., the distribution of Figure 10d leads to roughly similar response), in contrast to the calculations which capture the continuum limit. Figure 12a shows the slip distributions at 5 yr intervals; the

range of the slip axis corresponds to 300 yr at 35 mm/yr. Figure 12b shows distributions at 1 yr intervals for a range of the slip axis corresponding to 110 yr. This case involves an inherently discrete model and, consequently, the results are very different from those of cases I to V. Figures 12a and 12b show features like the cellular model discussed earlier with static/kinetic friction and $L = 0$. We note that the results of Figure 11e, for the strongly heterogeneous but continuum model of case V, suggest a way in which the two classes of models might be reconciled.

For the small $L$ values used in the calculations leading to Figures 11 and 12, essentially all slip over the velocity weakening range occurs as rupture instabilities. Figure 13 shows the average slip over 4 to 13 km depth as a function of time for case VI (oversized cells); the results are qualitatively similar to those shown in Figure 5 for the cellular static/kinetic friction model. Modest bursts in seismicity occur sometimes but not always before the large ruptures, and not every burst in seismicity is followed by a large rupture. Figures 12a, 12 b and 11e suggest that a brittle zone at the border of an aseismically slipping region can be a candidate for sequences of repeated small similar seismic events. The absence of small events in Figures 11a to 11d emphasizes the necessity for strong heterogeneity, adequate to stop seismic ruptures, for the occurrence of such sequences.

## DISCUSSION

Variations of fluid pressure in fault zone rocks have long been recognized as being important for fault dynamics [e.g., Nur, 1972; Sibson et al., 1975; Rice, 1979; Rudnicki, 1986]. Recent field, laboratory, and theoretical works [Byerlee, 1990, 1993; Blanpied et al. 1992; Sleep and Blanpied, 1992; Sibson, 1992; Rice 1992b] generated renewed interest in the effects of fluid pressure on frictional properties of faults. In this paper we simulate fault instabilities in models emphasizing the dependency of fault friction on complex distributions of geometrical irregularities, material heterogeneities and pore pressure variations. The simulations are done using two separate classes of models. The first class, taken as an approximate representation of geometric fault zone disorder, involves inherently discrete systems. This is modeled using both a cellular fault governed by static/kinetic friction [Ben-Zion and Rice, 1993], and a fault obeying rate- and state-dependent friction [Rice, 1993], but having numerical cells much larger than a nucleation size for slip instabilities. The second class involves a continuum system representing a smooth fault surface subjected to various degrees of pore pressure variations. This is modeled by a fault satisfying rate- and state-dependent constitutive law, using numerical cells much smaller than the critical nucleation size. In both classes of models, stress transfers due to slip episodes are governed by 3D continuum elasticity.

The results indicate that gradual pore pressure variations cannot explain the observed spatio-temporal complexity of seismic slip. It is possible, however, that the formation of seals [Byerlee, 1993] and localized zones that compact toward lithostatic $p$ [Sleep and Blanpied, 1992] could create strong and abrupt pore pressure fluctuations (e.g., of the type shown in Figures 2c and 10d) leading to complex seismic response. The simulations show that strong model heterogeneity of the form envisioned in our models with discrete (quasi-independent) numerical cells, representing geometric disorder in a fault system or abrupt property heterogeneity, is required for the generation of slip complexities, rough rupture fronts capable of radiating high frequency seismic waves, and FS statistics of earthquakes having a wide range of event sizes. The large earthquakes simulated by our discrete models are quasi-periodic; however, the character of the large events varies greatly from one sequence to the next. This is compatible with recent inversions of geodetic data along the central SAF, indicating that the 1934 and 1966 Parkfield earthquakes had substantially different rupture histories and/or distributions of slip [Segall and Du, 1993]. It thus appears that expectations for regular repeating sequences of earthquakes are unrealistic.

The frequency-size distribution of earthquakes is an important topic for both theoretical studies of earthquake dynamics and practical considerations of seismic hazard assessments. It is controversial whether FS statistics follow, in principle, a power law Gutenberg-Richter distribution, implying self-similarity over a broad range of earthquake magnitude, or whether earthquake statistics are characterized, intrinsically, by systematic strong deviations from self-similarity. The FS statistics of earthquakes observed along a given fault or in a seismogenic zone

with uniform width and one (dominant) mode of faulting show in many cases [e.g., Singh et al., 1983; Schwartz and Coppersmith, 1984; Main and Burton, 1984; Davison and Scholz, 1985; Main, 1987; Main and Burton, 1989; Trifu and Radulian, 1991; Wesnousky, 1993] strong deviations from self-similarity and local maxima in the rate of events having "characteristic" rupture dimensions. On the other hand, the FS statistics of observed earthquakes "sampling" irregular seismogenic zones such as southern California [Hileman et al., 1973] or the entire earth [Pacheco et al., 1992; Kagan, 1993] show a broad power law distribution. The observations suggest the existence of a "fundamental" FS distribution of earthquakes with features (and information) not accounted for by the simple GR relation, and that the latter broad power law GR statistics result from a secondary "averaging" process. This hypothesis is supported by the analysis of Ben-Zion and Rice [1993] and the present work.

Our results indicate that in a system characterized by geometric disorder (i.e., strength heterogeneity) spanning a narrow range of size scales, with stress transfer governed by 3D continuum elasticity, the FS statistics of earthquakes can be self-similar only over a narrow range of events smaller than a critical size. This is due to the fact that the stress concentrated in a 3D elastic solid at the edge of an expanding rupture grows with the rupture size. Thus, when the fault is characterized by a narrow range of geometric disorder, as in cases 1-4 of the analysis, the scaling of stress concentrations with the rupture size introduces a critical event size terminating the range of self-similar earthquake statistics. In such systems, events reaching the critical size become (on the average) unstoppable and they continue to grow to a size limited by a characteristic system dimension (see Figures 6 and 7). When, however, the system is characterized by a broad spectrum of geometric disorder, as in case 5 of the analysis, a corresponding broad spectrum of critical event sizes exists, and the phenomena discussed above are suppressed (Figure 8). The simulations indicate that local maxima in observed FS statistics correspond to dimensions of coherent brittle zones, such as the width of the seismogenic layer or the length of a fault segment bounded by barriers. Thus, observed FS statistics can be used to derive information on crustal thickness and fault zone structure. Our analysis offers a coherent explanation for both the GR statistics and the characteristic earthquake distribution of seismicity. The results suggest that the frequency of occurrence of moderate and strong earthquakes is enhanced with respect to self-similar distributions defined by small events, whether local maxima are observed in the FS statistics, or are averaged out.

The notion that earthquake are dynamically self-similar is very attractive on grounds of simplicity and analogy to other natural systems. However, the observations and analysis mentioned above indicate that continuum elasticity, coupled with a given segmented (discrete) fault zone structure, can produce a scale length (event diameter about $16h$ in our simulations; Ben-Zion and Rice [1993, Table 1]) terminating the self-similar regime of earthquake dynamics. The range of the self-similar regime can be extended by the existence of barriers that can stop ruptures at yet larger scales (see, e.g., Figures 2c and 8). As discussed in Ben-Zion and Rice [1993], a lower scale length that may limit self-similarity of earthquakes is the width of fault zones. This size scale may provide a basis for choosing $h$ when using the inherently discrete fault models. Analytic waveform fits of fault zone seismic trapped and head waves [Leary and Ben-Zion, 1992; Hough et al., 1993; Ben-Zion, 1993] indicate that the width of mature fault zones is of the order of a few hundreds of meters. Numerous seismological observations for small earthquakes, including the constancy of their radiated corner frequency [e.g., Chouet et al., 1978; Fletcher, 1980], the strong variation of their stress drops [e.g., Sacks and Rydelek, 1992; Hough et al., 1992] and their deviations from self-similar FS statistics [e.g., Aki, 1987; Malin et al., 1989; Rydelek and Sacks, 1989; Sacks and Rydelek, 1992], may be related to the width of the corresponding fault zone, and hence may be used to estimate the size of $h$. A third scale length that is an obvious candidate for disrupting self-similar seismic response is the width of the brittle seismogenic zone (about 10-15 km for the SAF; 30-100 km for subduction zones). Pacheco et al. [1992], using least-square analysis, claimed to have identified a break in the slope (power) of observed FS statistics of global earthquakes related to an (average) size of the seismogenic zone. However, Kagan [written communication, 1992; 1993, Figure 1], using the presumably superior maximum likelihood procedure, found that the gamma distribution (which does not show a break of slope

corresponding to the seismogenic layer) provides a good fit to the global FS data of the Harvard catalog.

It is worthwhile to close the discussion by mentioning a few shortcomings of our work. The simulations of Ben-Zion and Rice [1993] and the first section of our analysis model all slip episodes as brittle failures. This ignores time-dependent effects due to viscoelasticity in the lower crust [e.g., Li and Rice, 1987; Ben-Zion et al., 1993], and aseismic slip in creeping fault sections and in small slip patches within the computational grid [e.g., Dieterich, 1986; Tse and Rice, 1986; Rice, 1993]. As shown in Ben-Zion et al. [1993], however, slip velocities in the lower crust, and hence loading rates on the upper seismogenic layer, are expected to vary in time throughout an entire great earthquake cycle. Shortly after an 1857-type earthquake slip rates in the viscoelastic lower crust are higher than the far field plate velocity, while later in the cycle they are lower. High loading rates early in the cycle may result in clustering of earthquakes in space and time, while low loading rates late in the cycle may result in an overall decrease of activity. Similarly, time-varying stress concentrations due to aseismic slip in small slip patches (i.e., patches with size $h < h^*$) and in creeping fault sections can have important effects on foreshock-mainshock-aftershock sequences. The modeling of these time-dependent effects can be achieved by incorporating features of rate- and state- dependent friction into the constitutive stress-slip relation.

The quasi-static simulations of Rice [1993], Ben-Zion and Rice [1993] and the present work model inertial effects during seismic slip only very approximately (e.g., by a seismic radiation damping term for rapid slip along the fault, or by allowance for dynamic overshoot in rupture arrest). Thus, calculated final states of stress on the fault after failures can be correct only in an approximate sense. It is important to examine whether a more precise treatment of elastodynamics during instabilities would qualitatively modify critical features of the modeling results, in a way that would change our conclusions on features responsible for slip complexity. As suggested by Madariaga and Cochard [1992], it is possible that the wave-mediated arrest of a dynamic slip event will leave a heterogeneous distribution of residual static stress on the ruptured surface. Such heterogeneity, if strong enough, could affect the nucleation and arrest locations of subsequent ruptures and may become a mechanism for sustaining complex seismic behavior. Also, the large off-fault stresses generated near a rupture front as it accelerates toward its limiting speed [e.g., Rice, 1980] could promote non-planar rupture features and be a further factor inducing complex response.

## ACKNOWLEDGEMENTS

We thank Joe Andrews, Yan Kagan, Leon Knopoff and Raul Madariaga for discussions, and Steve Brown for a program generating fractal surfaces. The work was supported by the Southern California Earthquake Center (subcontract 569928 from USC, based on NSF support), and the USGS National Earthquake Hazard Reduction Program (grant 1434-93-G-2276).

## REFERENCES

Aki, K., Magnitude-frequency relation for small earthquakes: A clue to the origin of $f_{max}$ of large earthquakes, *J. Geophys. Res.*, **92**, 1349 - 1355, 1987.

Andrews, D. J., A stochastic fault model 1. static case, J. *Geophys. Res.*, **85**, 3867-3877, 1980.

Andrews, D. J., A stochastic fault model 2. time-dependent case, *J. Geophys. Res.*, **86**, 10821-10834, 1981.

Antolik, M., W. Foxall, A. Michelini and T. E. McEvilly, Microearthquake clusters at Parkfield: Constraint on fault-zone structure and failure processes, *EOS Trans. Amer. Geophys. Union*, **72**, 483, 1991.

Aviles, C. A., C. H. Scholz, and J. Boatwright, Fractal analysis applied to characteristic segments of the San Andreas Fault, *J. Geophys. Res.*, **92**, 331 - 334, 1987.

Bak, P. and C. Tang, Earthquakes as self-organized critical phenomena, *J. Geophys. Res.*, **94**, 15635 - 15637, 1989.

Ben-Zion, Y., Structure of a few fault segments in California from travel time inversion and waveform modeling of fault zone seismic waves, *Ann. Geophys.*, **11**, suppl. I, c47, 1993.
Ben-Zion, Y. and J. R. Rice, Earthquake failure sequences along a cellular fault zone in a 3D elastic solid containing asperity and nonasperity regions, *J. Geophys. Res.*, **98**, 14109-14131, 1993.
Ben-Zion, Y., J. R. Rice and R. Dmowska, "Interaction of the San Andreas Fault creeping segment with adjacent great rupture zones and earthquake recurrence at Parkfield", *J. Geophys. Res.*, **98**, 2135-2144, 1993
Blanpied, M. L., D. A. Lockner and J. D. Byerlee, Fault stability inferred from granite sliding experiments at hydrothermal conditions, *Geophys. Res. Letters*, **18** (4), 609-612, 1991.
Blanpied, M. L., D. A. Lockner and J. D. Byerlee, An earthquake mechanism based on rapid sealing of faults, *Nature*, **358**, 574-358, 1992.
Brown, S., Simple mathematical models of a rough fracture, manuscript, 1993.
Brown, S., J. B. Rundle and C. H. Scholz, A simplified spring-block model of earthquakes, *Geophys. Res. Lett.*, **18**, 215 - 218, 1991.
Burridge, R. and L. Knopoff, Model and theoretical seismicity, *Bull. Seismol. Soc. Amer.*, **57**, 341 - 371, 1967
Byerlee, J. D. Friction, overpressure and fault normal compression, *Geophys. Res. Lett.*, **17**, 2109-2112, 1990.
Byerlee, J. Model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology* , **21**, 303-306, 1993.
Carlson, J. M. and J. S. Langer, Mechanical model of an earthquake, *Physical Review A*,, **40**, 6470 - 6484, 1989.
Carlson, J. M., J. S. Langer, B. Shaw and C. Tang, Intrinsic properties of a Burridge-Knopoff model of a fault, *Physical Review A*, **44**, 884 - 897, 1991.
Chinnery, M., The stress changes that accompany strike-slip faulting, *Bull. Seismol. Soc. Amer.*, **53**, 921-932, 1963.
Chouet, B., K. Aki, and M. Tsujiura, Regional variation of the scaling law of earthquake source spectra, *Bull. Seismol. Soc. Amer.*, **68**, 49 - 79, 1978.
Christensen, K. and Z. Olami, Variation of the Gutenberg-Richter b values and nontrivial temporal correlations in a spring-block model for earthquakes, *J. Geophys. Res.*, **97**, 8729 - 8735, 1992.
Davison, F. C. and C. H. Scholz, Frequency-moment distribution of earthquakes in the Aleutian arc: a test of the characteristic earthquake model, *Bull. Seismol. Soc. Amer.*, **75**, 1349-1361, 1985.
Dieterich, J. H., A model for the nucleation of earthquake slip, *in Earthquake Source Mechanics, Geophys. Monogr. Ser.*, vol. **37**, edited by S. Das, J. Boatwright, and C. H. Scholz, pp. 37-47, AGU, Washington D. C., 1986.
Fletcher, J. B., Spectra from high-dynamic range digital recording of Oroville, California, aftershocks and their source parameters, *Bull. Seismol. Soc. Amer.*, **70**, 735 - 755, 1980.
Frankel, A., High-frequency spectral falloff of earthquakes, fractal dimension of complex rupture, *b* value, and the scaling of strength on faults, *J. Geophys. Res.*, **96**, 6291 - 6302, 1991.
Gao, H., J. J. Rice, and J. Lee, Penetration of a quasi-statically slipping crack into a seismogenic zone of heterogeneous fracture resistance, *J. Geophys. Res.*, **96**, 21,535 - 21,548, 1991.
Hanks, T. C., and H. Kanamori, A moment magnitude scale, *J. Geophys. Res.*, **84**, 2348 - 2350, 1979.
Harris, R. A., and P. Segall, Detection of a locked zone at depth on the Parkfield, California, segment of the San Andreas fault, *J. Geophys. Res.*, **92**, 7945 - 7962, 1987.
Hileman, J. A., C. R. Allen and J. M. Nordquist, Seismicity of the southern California region 1st Jan. 1932 to 31st Dec. 1972, *Contr. Div. Geol. Planet. Sci., Calif. Inst. Techn., NO. 2385*, 1973.
Horowitz, F. G. and A. Ruina, Slip patterns in a spatially homogeneous fault model, *J. Geophys. Res.*, **94**, 10279-10298, 1989.
Hough, S. E., H. Gue, A. Lerner-Lam, S. Seeber and C. Scholz, Stress drop scaling: results

from empirical Green's function study of Loma Prieta aftershocks, *Seismological Research Letters*, **63**(1), p. 24, 1992.
Hough, S. E., Y. Ben-Zion and P. Leary, Tectonic structure of the southern end of the 4/22/92 Joshua Tree earthquake rupture zone from analysis of fault zone waves, submitted to *Bull. Seismol. Soc. Amer.*, 1993.
Ito, K. and M. Matsuzaki, Earthquakes as self organized critical phenomena, *J. Geophys. Res.*, **95**, 6853 - 6860, 1990.
Kagan, Y. Y., Stochastic model of earthquake fault geometry, *Geophys. J. R. Astr. Soc.*, **71**, 659-691, 1982.
Kagan, Y. Y., Fractal dimension of brittle fracture, *J. Nonlinear Sci.*, **1**, 1 - 16, 1991.
Kagan, Y. Y., Seismicity: turbulence of solids, *Nonlinear Sci. Today*, 2, 1-13, 1992.
Kagan, Y. Y., Statistics of characteristic earthquakes, *Bull. Seismol. Soc. Amer.*, **83**, 7-24, 1993.
Kagan, Y. Y. and L. Knopoff, Stochastic synthesis of earthquake catalogs, *J. Geophys. Res.*, **86**, 2853-2862, 1981.
Leary P., Deep borehole log evidence for fractal distribution of fractures in crystalline rock, *Geophys. J. Int.*, **107**, 615 - 627, 1991.
Leary, P. and Y. Ben-Zion, A 200 m wide fault zone low velocity layer on the San Andreas fault at Parkfield: Results from analytic waveform fits to trapped wave groups, *Seismol. Res. Lett.*, **63**(1), 62, 1992.
Li, V. C., and J. R. Rice, Crustal deformation in great California earthquake cycles, *J. Geophys. Res.*, **92**, 11533-1151, 1987.
Madariaga, R. and A. Cochard, Heterogeneous faulting and friction, Proc. Int. Symp. on Earthquake Disaster Prevention, Mexico City, ed. R. Melli, vol. I, pp. 103-118, CENAPRED, 1992
Main, I. G., A characteristic earthquake model of the seismicity preceding the eruption of Mount St. Helens on 18 May 1980, *Physics of the Earth and Planetary Interiors*, **49**, 283-293, 1987.
Main, I. G. and P. W. Burton, Physical links between crustal deformation, seismic moment and seismic hazard for regions of varying seismicity, *Geophys. J. R. Astr. Soc*, **79**, 469-488, 1984.
Main, I. G. and P. W. Burton, Seismotectonics and the earthquake frequency-magnitude distribution in the Aegean area, *Geophys. J.*, **98**, 575-586, 1989.
Malin, P. E., S. N. Blakeslee, M. G. Alvarez, and A. J. Martin, Microearthquake imaging of the Parkfield asperity, *Science*, **244**, 557 - 559, 1989.
Nur, A., Dilatancy, pore fluids, and premonitory variations of $t_s/t_p$ travel times, *Bull. Seismol. Soc. Amer.*, **62**, 1217-1222, 1972.
Okubo, P. and K. Aki, Fractal geometry in the San Andreas fault system, *J. Geophys. Res.*, **92**, 345 - 355, 1987.
O'Neill, M. E., Source dimensions and stress drops of small earthquakes near Parkfield, California *Bull. Seismol. Soc. Amer.*, **74**, 27 - 40, 1984.
Pacheco, J. F., C. H. Scholz, and L. R. Sykes, Changes in frequency-size relationship from small to large earthquakes, *Nature*, **255**, 71-73, 1992.
Power, W. L. and T. E. Tullis, Euclidian and fractal models for the description of rock surface roughness, *J. Geophys. Res.*, **96**, 415 - 424, 1991.
Power, W. L., T. E. Tullis and J. D. Weeks, Roughness and wear during brittle faulting, *J. Geophys. Res.*, **93**, 15268 - 15287, 1988.
Rice, J. R. Theory of precursory processes in the inception of earthquake rupture, *Gerlands Beitr. Geophysics*, **88**, 91-127, 1979.
Rice, J. R., The mechanics of earthquake rupture, in *Physics of the Earth's Interior*, ed. A. M. Dziewonski and E. Boschi, pp. 555-649, Italian Physical Society / North Holland, Amsterdam, 1980.
Rice, J. R., Variable earthquake recurrence intervals at Parkfield, Abstract, *EOS Trans. Amer. Geophys. Union*, **73**, 1992a.
Rice, J. R., Fault stress states, pore pressure distributions, and the weakness of the San Andreas

fault, *in Fault Mechanics and Transport Properties of Rocks*, 475-503, Academic, London, 1992b.
Rice, J. R., Spatio-temporal complexity of slip on a fault, *J. Geophys. Res.*, **98**, 9885-9907, 1993.
Rice, J. R. and Y. Ben-Zion, Prolong dynamic oscillatory crack motion due to asperity encounters, *EOS Trans. Amer. Geophys. Union*, **74**, 292, 1993.
Rudnicki, J. W., Slip on an impermeable fault in a fluid-saturated rock mass, *in Earthquake Source Mechanics, Geophys. Monogr. Ser.*, vol. **37**, edited by S. Das, J. Boatwright, and C. H. Scholz, pp. 81-89, AGU, Washington D. C., 1986.
Rundle, J. B., A physical model for earthquakes 1. Fluctuations and interactions, *J. Geophys. Res.*, **93**, 6237 - 6254, 1988.
Rydelek, P. A. and I. S. Sacks, Testing the completeness of earthquake catalogues and the hypothesis of self-similarity, *Nature*, **337**, 251 - 253, 1989.
Sacks, S. E. and P. A. Rydelek, Quantum earthquake concept, Abstract, *Seismological Research Letters*, **63**(1), p. 75, 1992.
Schwartz, D. P. and K. J. Coppersmith, Fault behavior and characteristic earthquakes: examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res.*, **89**, 5681-5698, 1984.
Segall, P. and Y. Du, How similar were the 1934 and 1966 Parkfield earthquakes?, *J. Geophys. Res.*, **98**, 4527-4538, 1993.
Sibson, R. H., Implications of fault-valve behavior for rupture nucleation and recurrence, *Tectonophysics.*, **211**, 283-293, 1992.
Sibson, R. H., J. McMoore and A. H. Rankine, Seismic pumping - A hydrothermal fluid transport mechanism, *J. Geol. Soc. London*, **131**, 653-659, 1975.
Singh, S. K., M. Rodriguez and L. Esteva, Statistics of small earthquakes and frequency of occurrence of large earthquakes along the Mexican subduction zone, *Bull. Seismol. Soc. Am.*, **73**, 1779-1796, 1983.
Sleep, N. H. and M. L. Blanpied, Creep, compaction and the weak rheology of major faults, *Nature*, **359**, 687-692, 1992.
Stuart, W. D. and G. M. Mavko, Earthquake instability on a strike-slip fault, *J. Geophys. Res.*, **84**, 2153-2160, 1979.
Stuart, W. D., Forecast model for large and great earthquakes in Southern California, *J. Geophys. Res.*, **91**, 13,771-13,786, 1986.
Stuart, W. D., R. J. Archuleta and A. G. Lindh, Forecast model for moderate earthquakes near Parkfield, California, *J. Geophys. Res.*, **90**, 592-604, 1985.
Trifu, C.-I., and M. Radulian, Frequency-magnitude distribution of earthquakes in Vrancea: relevance for a discrete model, *J. Geophys. Res.*, **96**, 4301-4311, 1991.
Tse, S. T. and J. R. Rice, Crustal earthquake instability in relation to the depth variation of frictional slip properties, *J. Geophys. Res.*, **91**, 9452-9472, 1986.
Wesnousky, S. G., The Gutenberg-Richter or characteristic earthquake distribution, which is it?, submitted to *Bull. Seismol. Soc. Amer.*, 1993.
Yamashita, T., Application of fracture mechanics to the simulation of seismicity and recurrence of characteristic earthquakes on a fault, *J. Geophys. Res.*, **98**, 12,019-12,032, 1993.
Yamashita, T. and L. Knopoff, Model for intermediate-term precursory clustering of earthquakes, *J. Geophys. Res.*, **97**, 19,873-19,879, 1992.

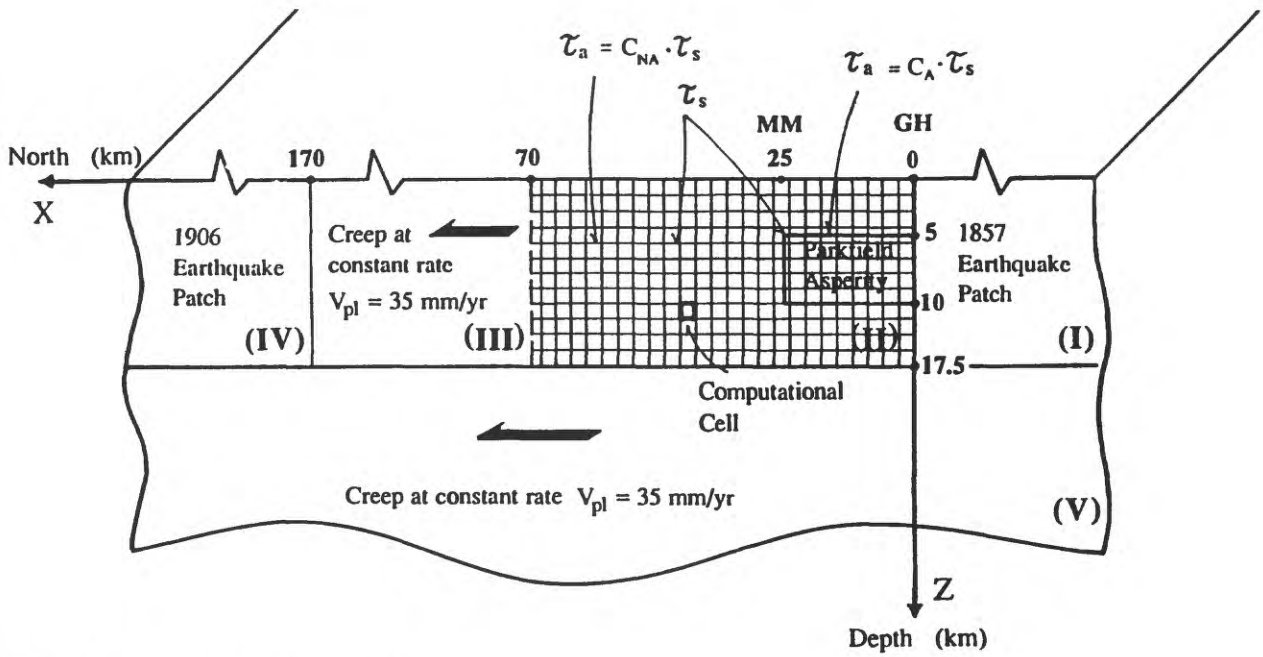



Figure 1. A model for the central San Andreas fault (SAF). $\tau_s$ is static strength assumed homogeneous over the computational region, $\tau_a$ is arrest stress, $C_A$ and $C_{NA} \geq C_A$ are numerical coefficients, GH and MM mark approximate positions of Gold Hill and Middle Mountain.

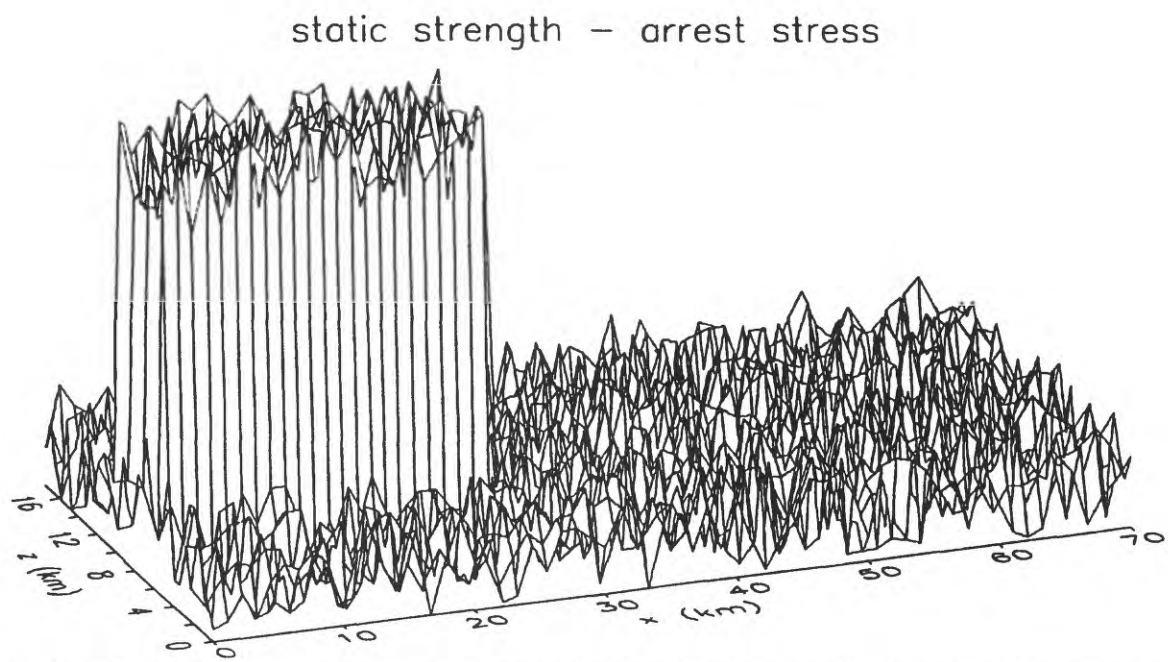



Figure 2a. Case 2 of text. Assumed stress drop distribution with an explicit "Parkfield" asperity and uncorrelated random property variations. In the asperity region the average arrest stress is $\tau_a = 0.1\tau_s$, while in the non-asperity region $\tau_a = 0.8\tau_s$. The amplitude of the random fluctuations is $0.1\tau_s$.

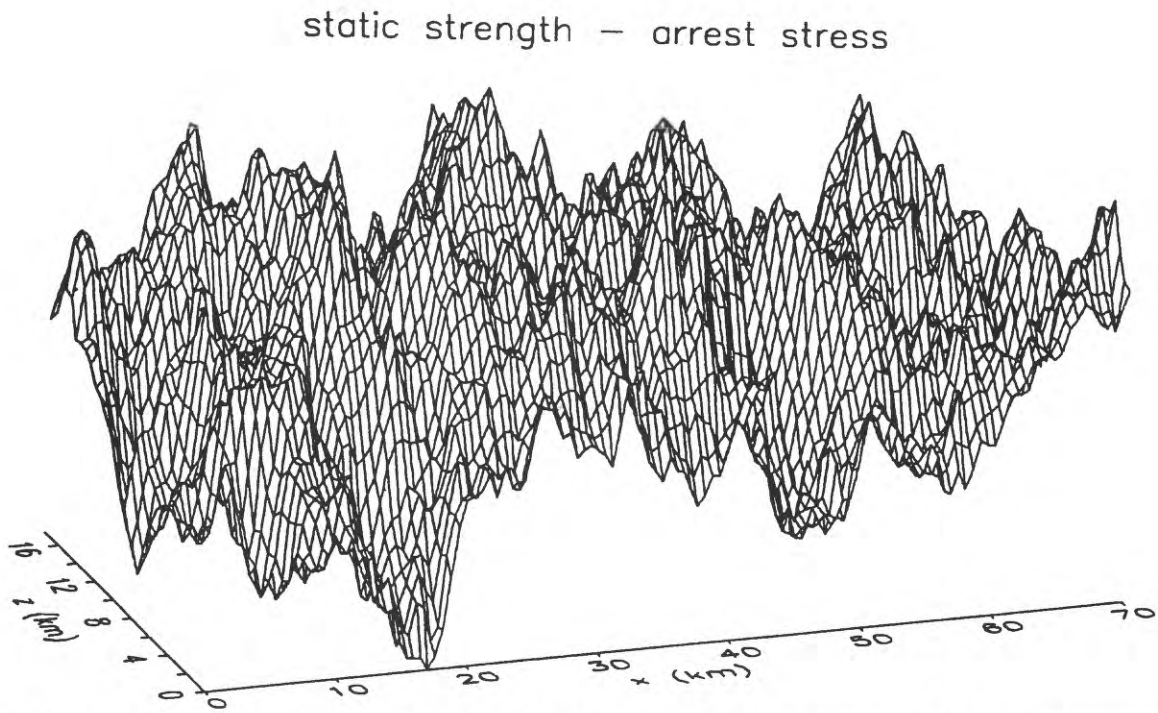



Figure 2b. Case 4 of text. Assumed stress drops having a fractal distribution characterized by fractal dimension 2.3, mean value 0.5, and standard deviation 0.2.

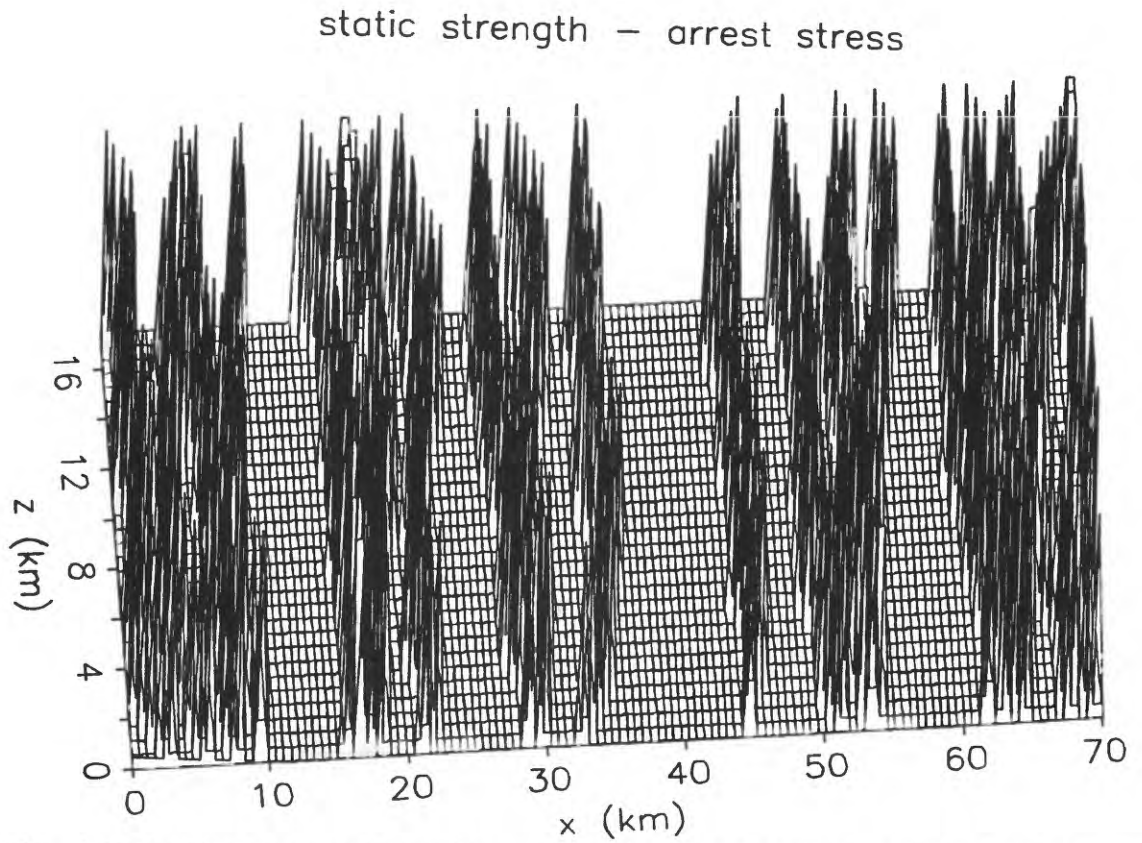



Figure 2c. Case 5 of text. Assumed stress drops representing "weak" fault segments separated by "strong" boundaries. See text for more explanation.

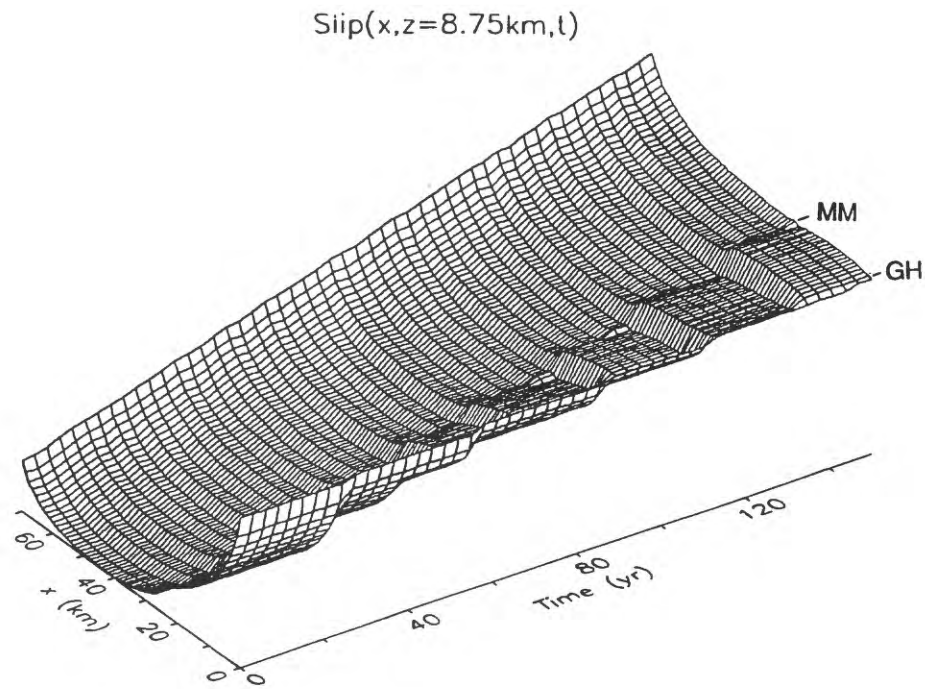



**Figure 3a. Slip (normalized to the range 0-1) at a depth of 8.75 km as a function of distance along the SAF and time for the property distribution of Figure 2a (case 2 of text).**

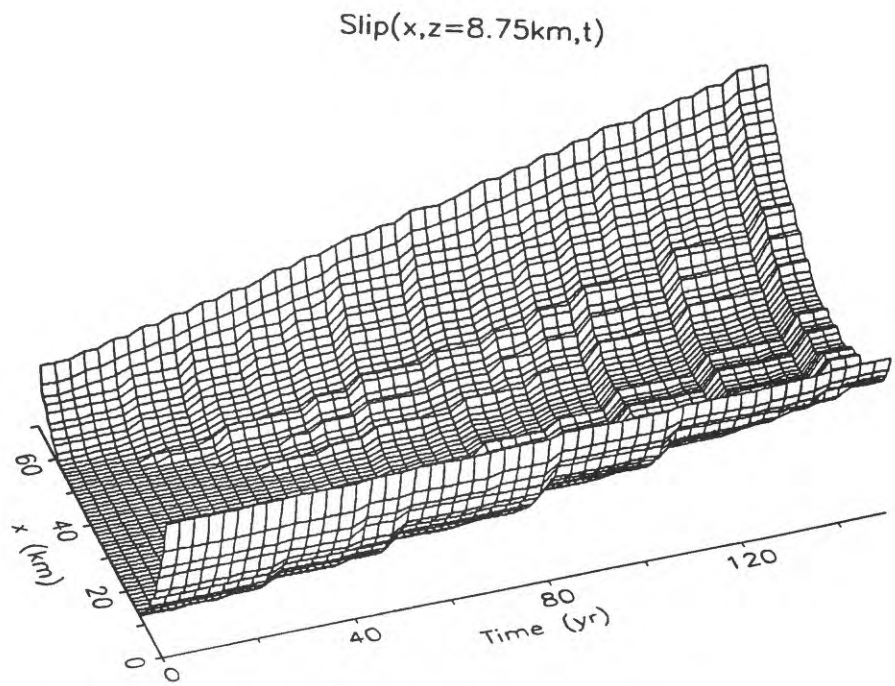



**Figure 3b. Same as (a) for the property distribution of Figure 2c (case 5 of text).**

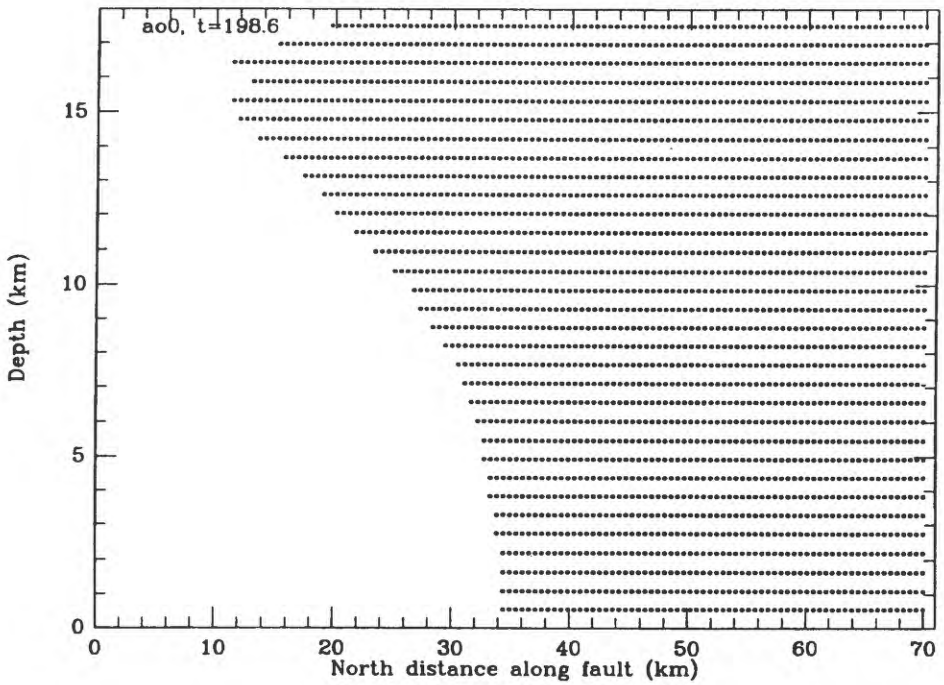



Figure 4a. Rupture area (dots) with a smooth front in simulation with constant fault properties (case 1 of text).

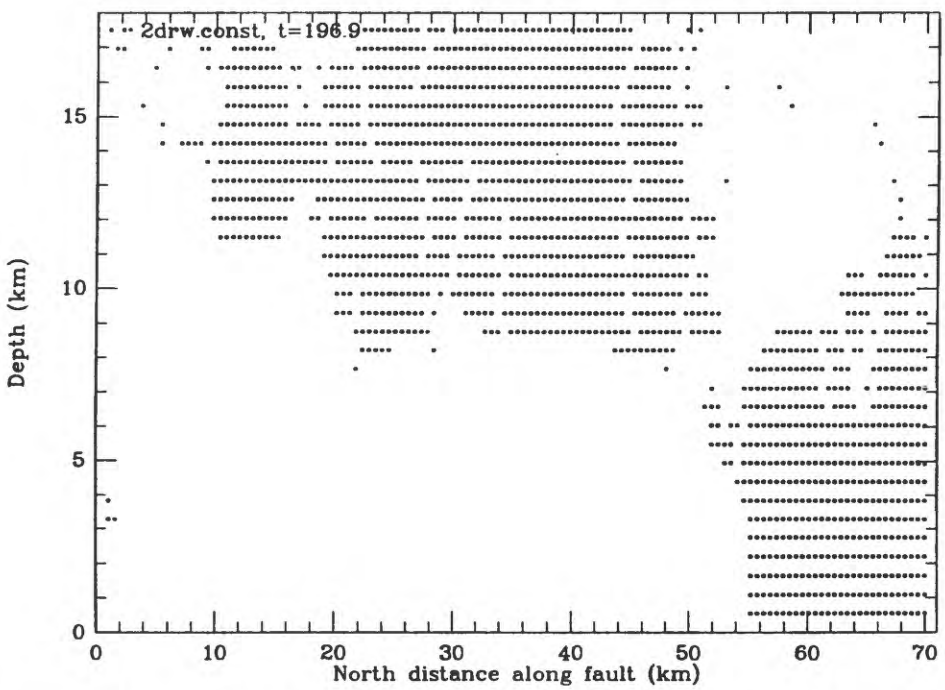



Figure 4b. Rupture area (dots) with a rough front in simulation with space-variable fault properties (case 5 of text, Figure 2c).

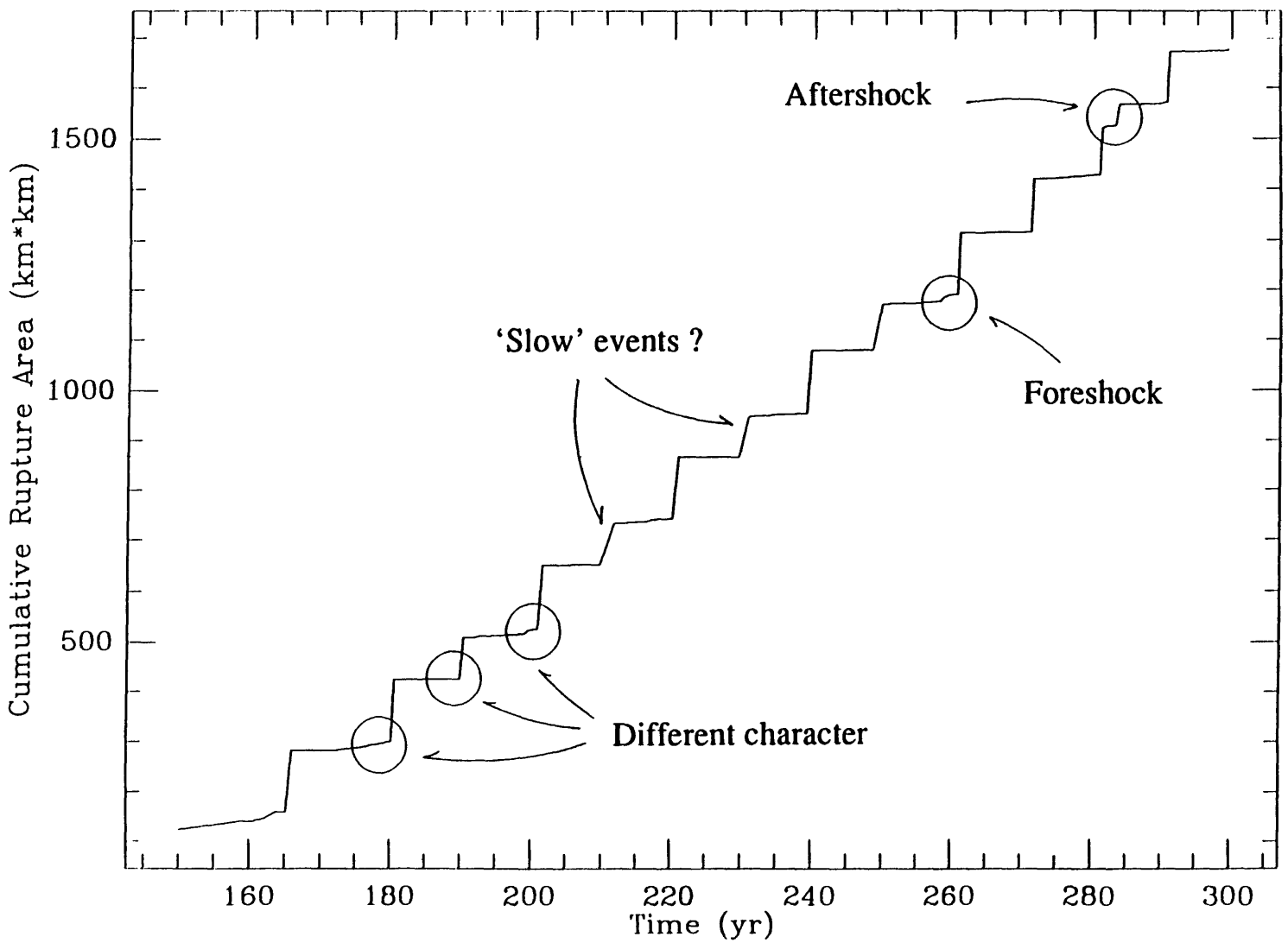



Figure 5. Cumulative rupture area at the Parkfield asperity region for case 1 of text. Note the diversity in the character of the large events.

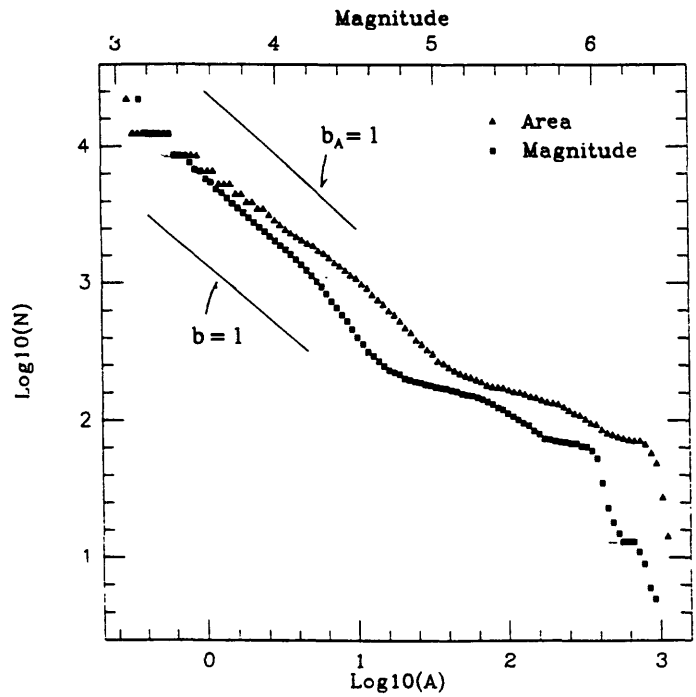



Figure 6a. Cumulative frequency-size (FS) statistics of earthquake magnitude (squares) and rupture area (triangles) during 150 analysis yr in the computational grid for the stress drop distribution case 1 of text (i.e., like Figure 2a but without random fluctuations). Units of rupture area are $km^2$.

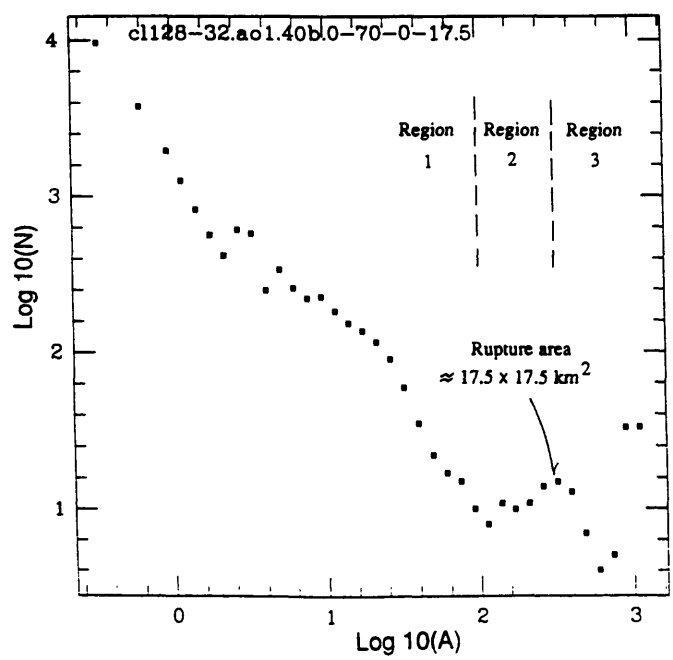



Figure 6b. Same as (a) for noncumulative FS statistics of rupture area.

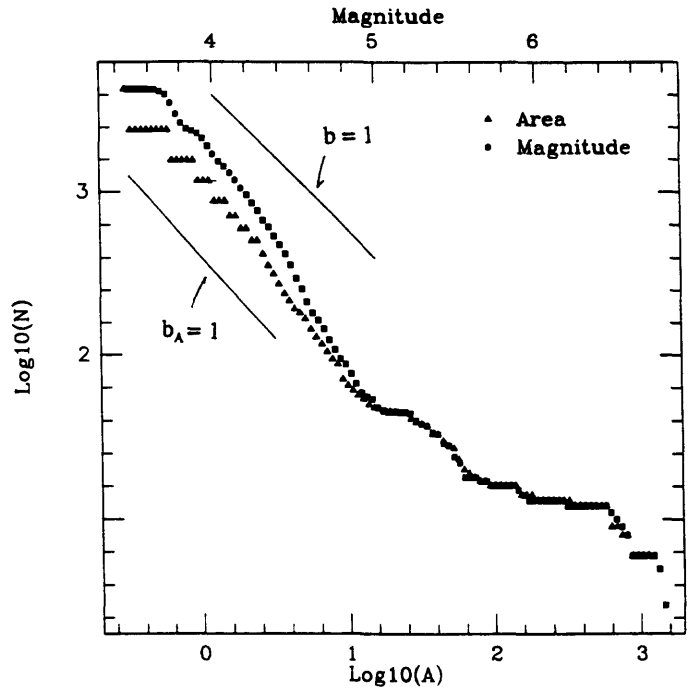



Figure 7a. Cumulative FS statistics of earthquake magnitude and rupture area during 150 analysis yr in the computational grid for the fractal stress drop distribution, case 4 of text (Figure 2b).

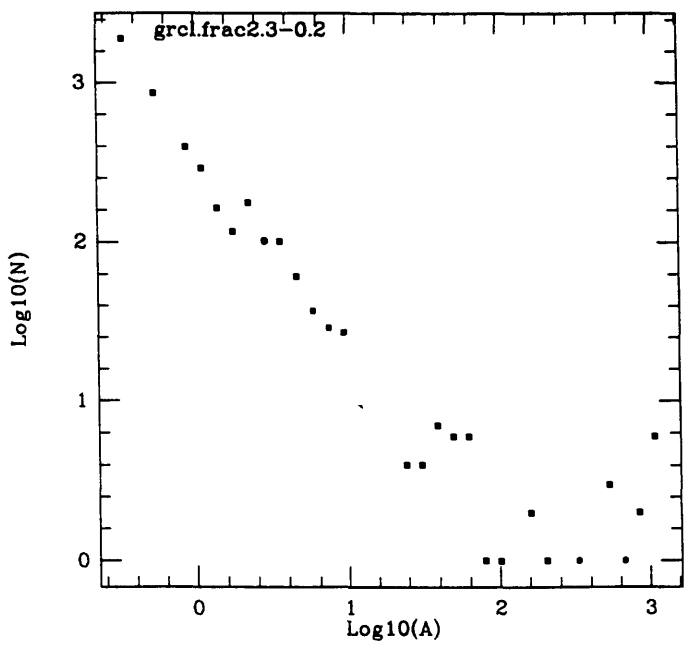



Figure 7b. Same as (a) for noncumulative FS statistics of rupture area.

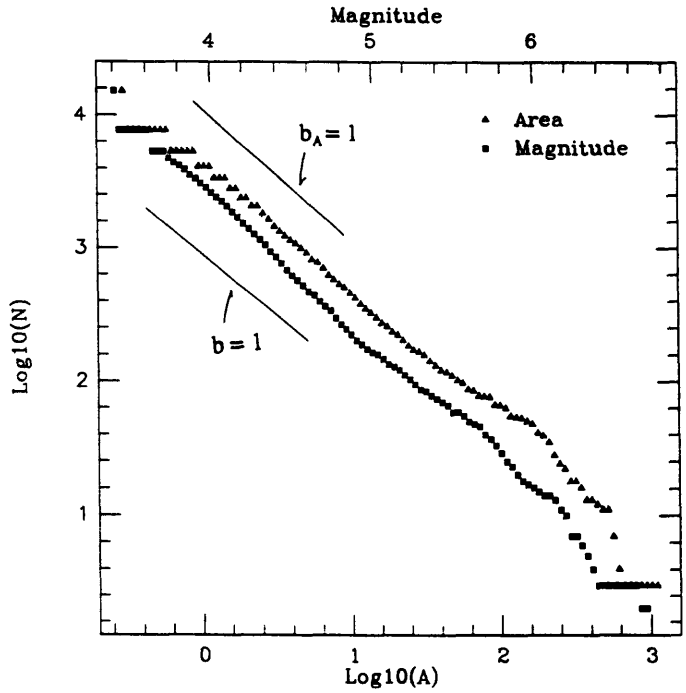



Figure 8a. Cumulative FS statistics of earthquake magnitude and rupture area during 150 analysis yr in the computational grid for the stress drop distribution representing weak segements of variable size with strong boundaries, case 5 of text (Figure 2c).

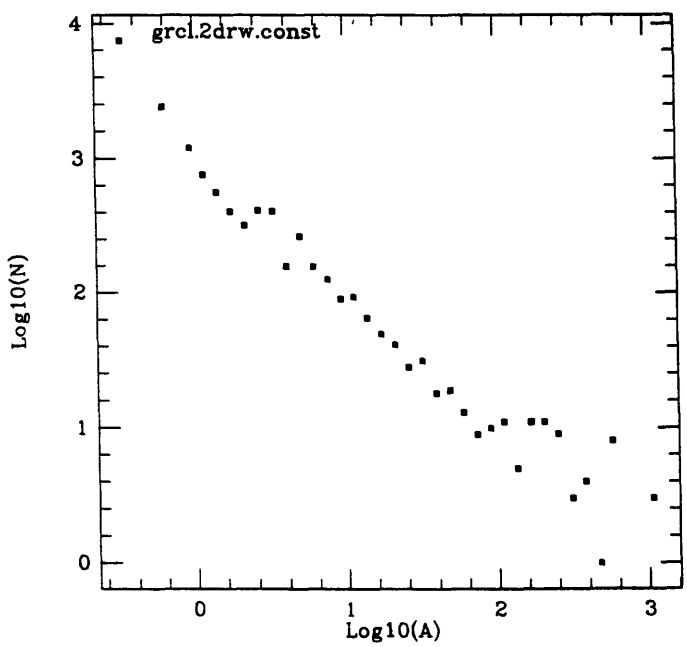



Figure 8b. Same as (a) for noncumulative FS statistics of rupture area.

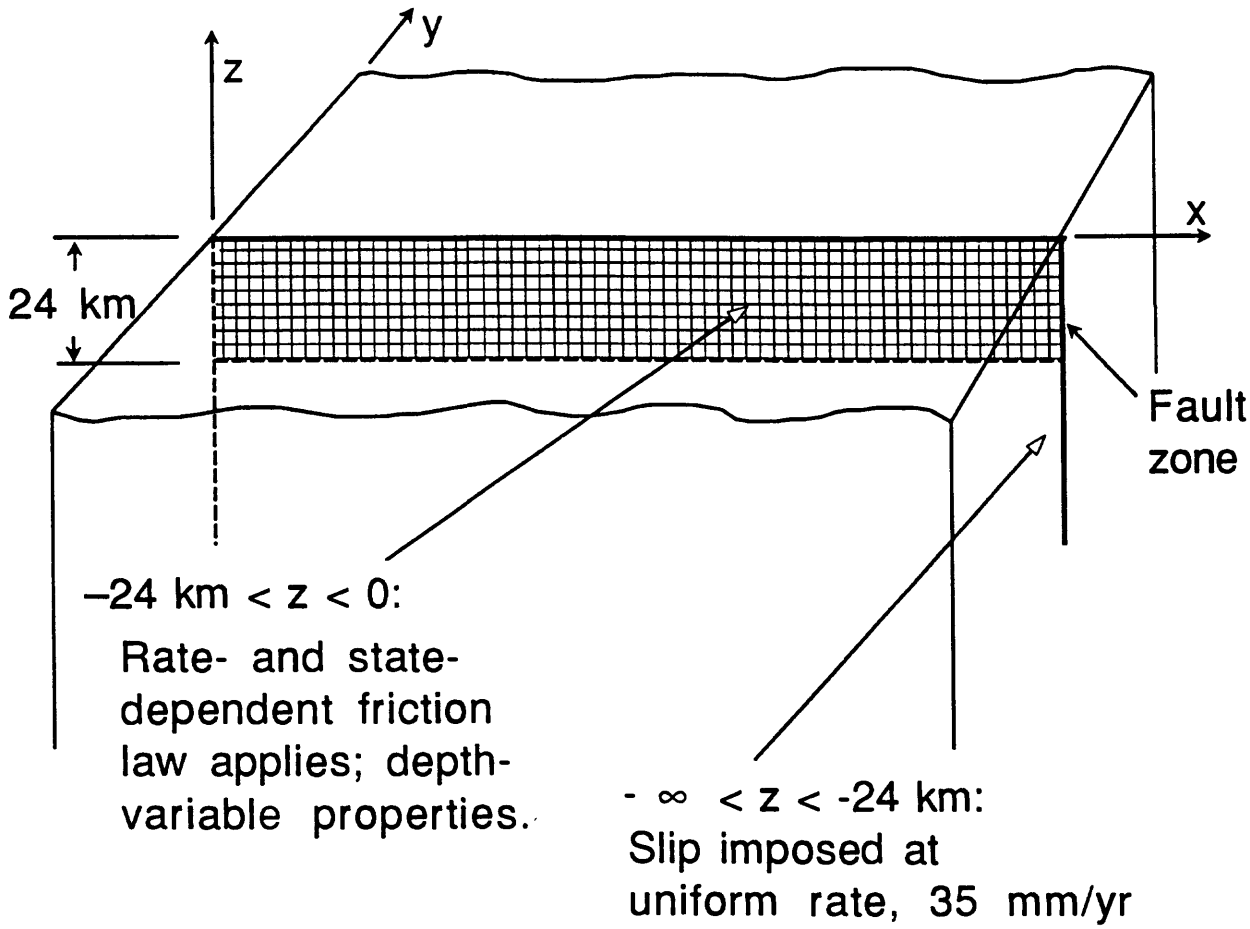



Figure 9. Elastic half-space with vertical half-plane strike slip fault. Slip imposed at 35 mm/yr below 24 km and calculated above 24 km according to rate- and state-dependent friction law, like in Rice [1993]. The cases considered in this work constrain slip to vary only with depth.

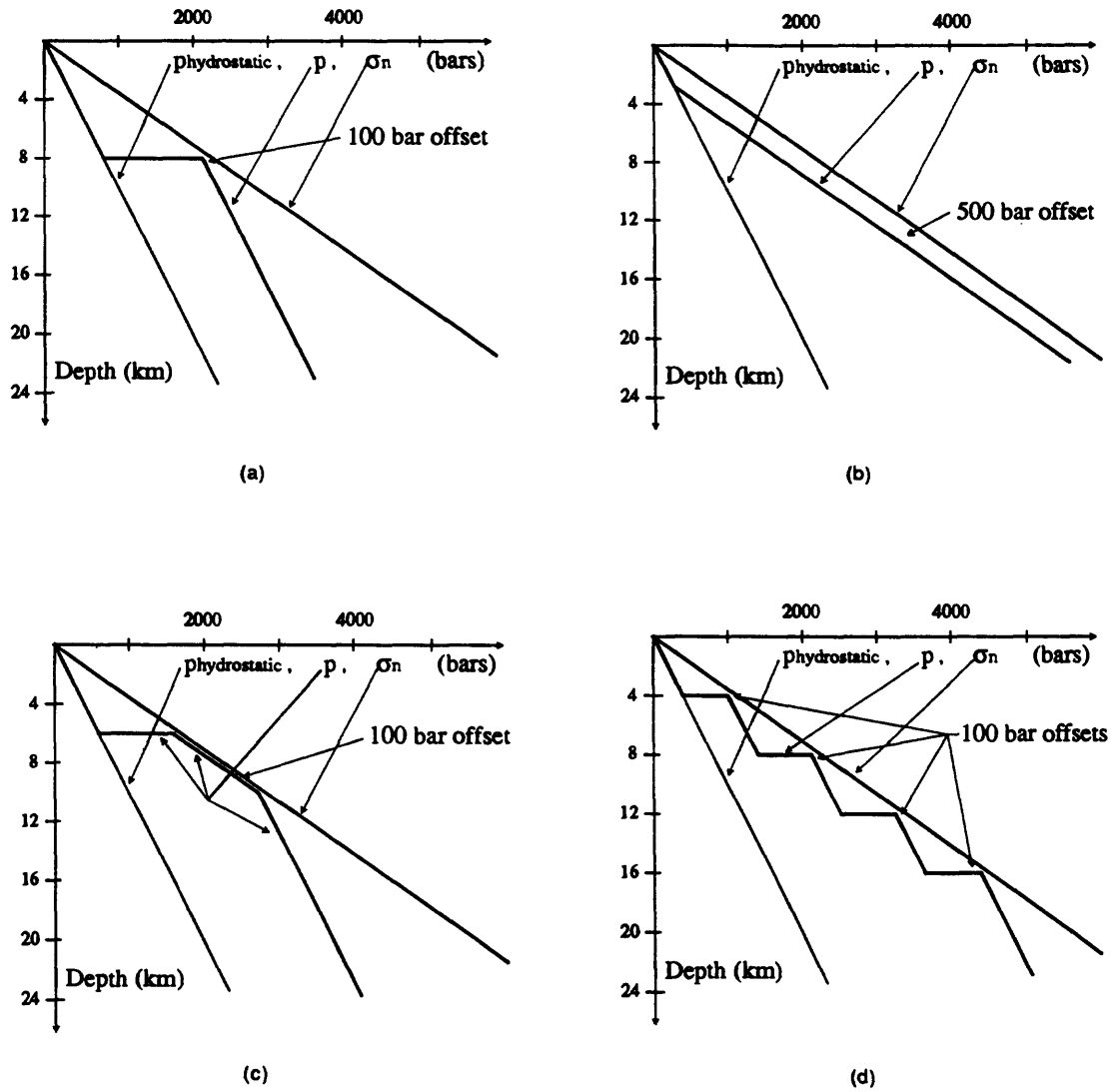



Figure 10. Pore pressure distributions used in calculations (Cases I to VI) leading to Figures 11 and 12. Thin lines in each panel show hydrostatic pore pressure $p$; this pressure is used in for Case I (Figure 11a). (a) Step to 10 MPa offset from lithostatic pressure $\sigma_n$ at 8 km; Case II (Figure 11b). (b) Lithostatic gradient with 50 MPa offset at depth; Case III (Figure 11c) and Case VI (Figure 12). (c) Step at 6 km with lithostatic gradient to 10 km and 10 MPa offset from $\sigma_n$; Case IV (Figure 11d). (d) Four steps, each to 10 MPa offset from $\sigma_n$, at 4, 8, 12, and 16 km; Case V (Figure 11e).

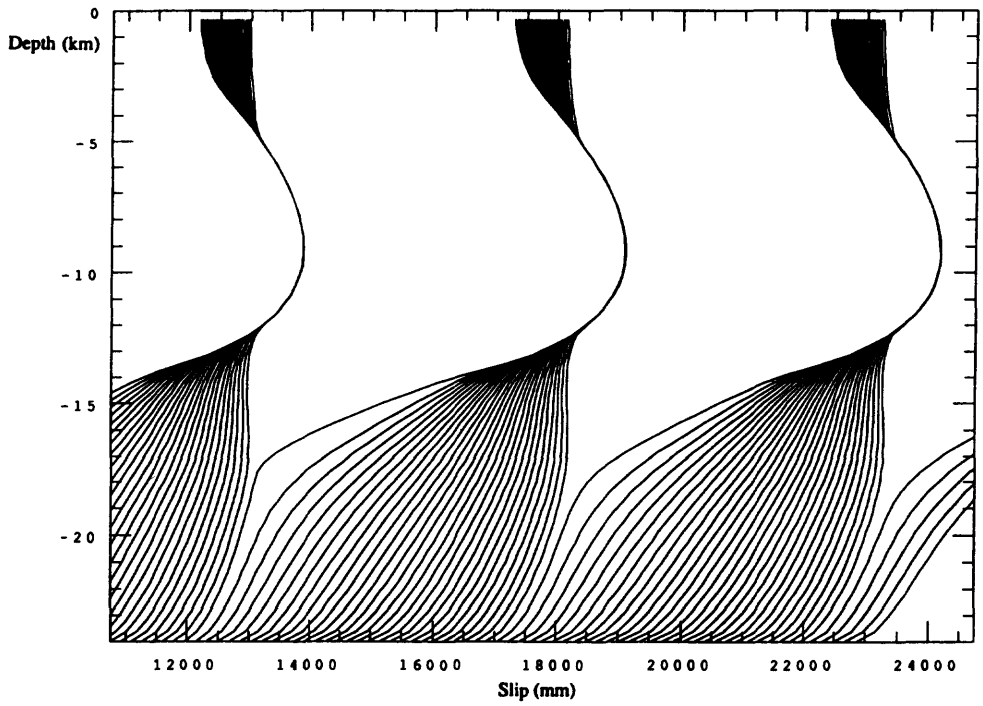



Figure 11a. Slip versus depth at 5 yr intervals for case I of text; hydrostatic pore pressure.

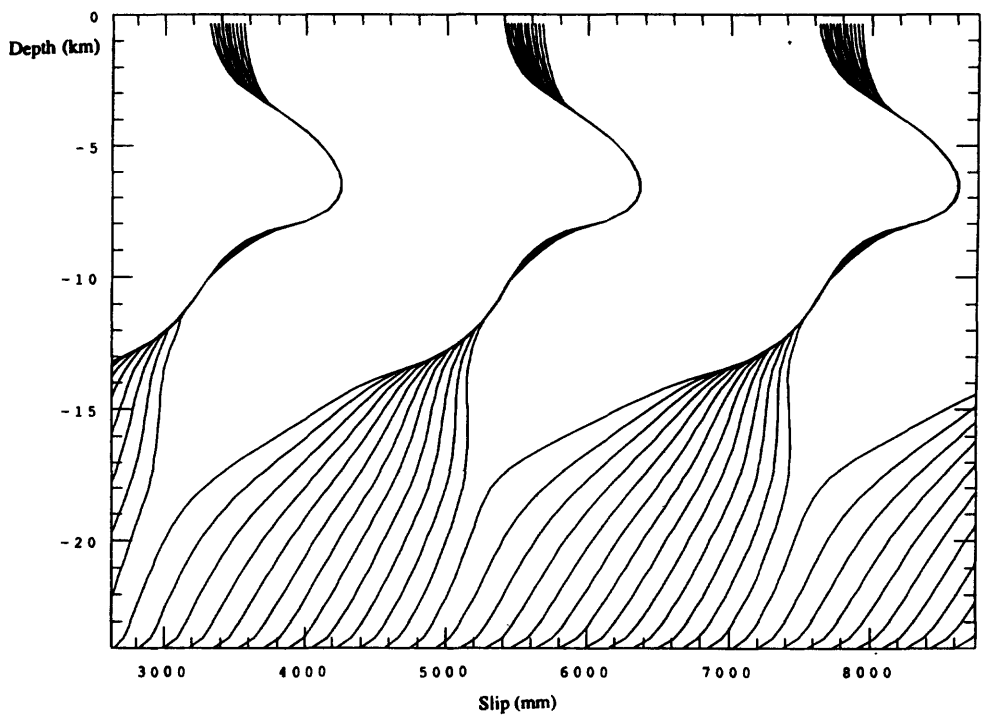



Figure 11b. Slip versus depth at 5 yr intervals for case II of text; step in $p$ to 10 MPa offset from $\sigma_n$ at 8 km (Figure 10a).

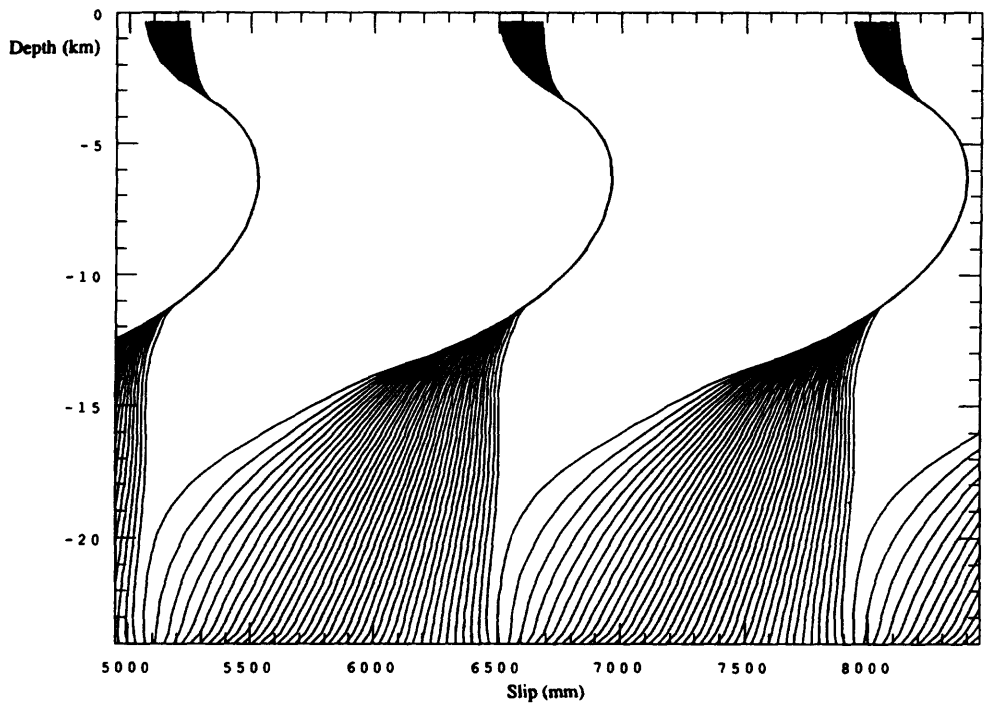



Figure 11c. Slip versus depth at 1 yr intervals for case III of text; $p$ has lithostatic gradient at depth and is offset 50 MPa from $\sigma_n$ (Figure 10b).

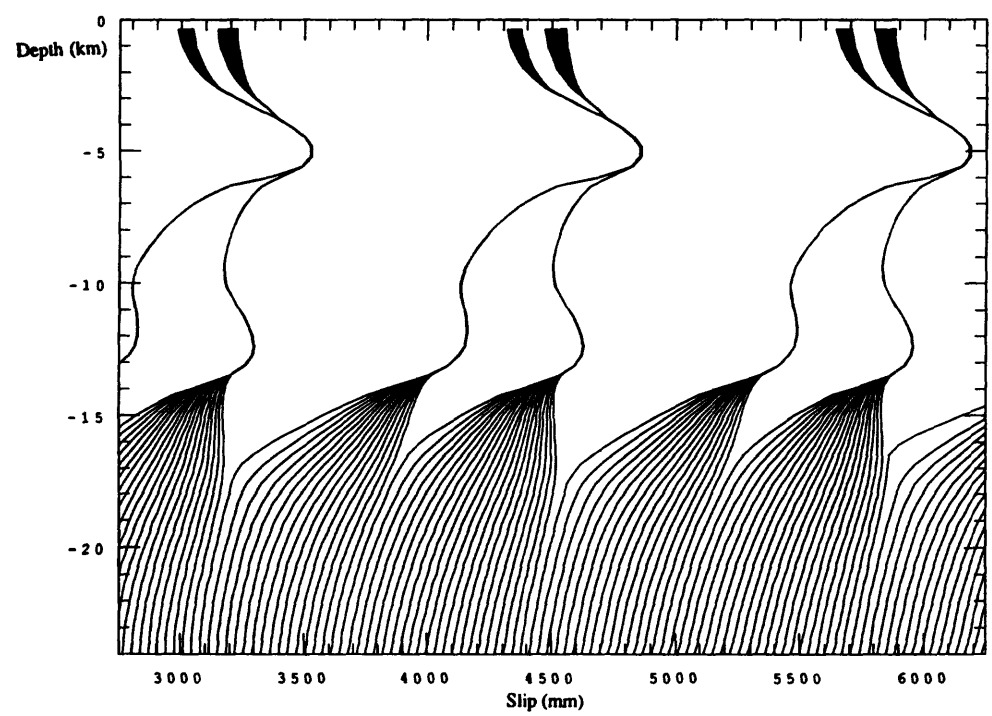



Figure 11d. Slip versus depth at 1 yr intervals for case IV of text; $p$ has step to 10 MPa offset from $\sigma_n$ at 6 km and lithostatic gradient to 10 km (Figure 10c).

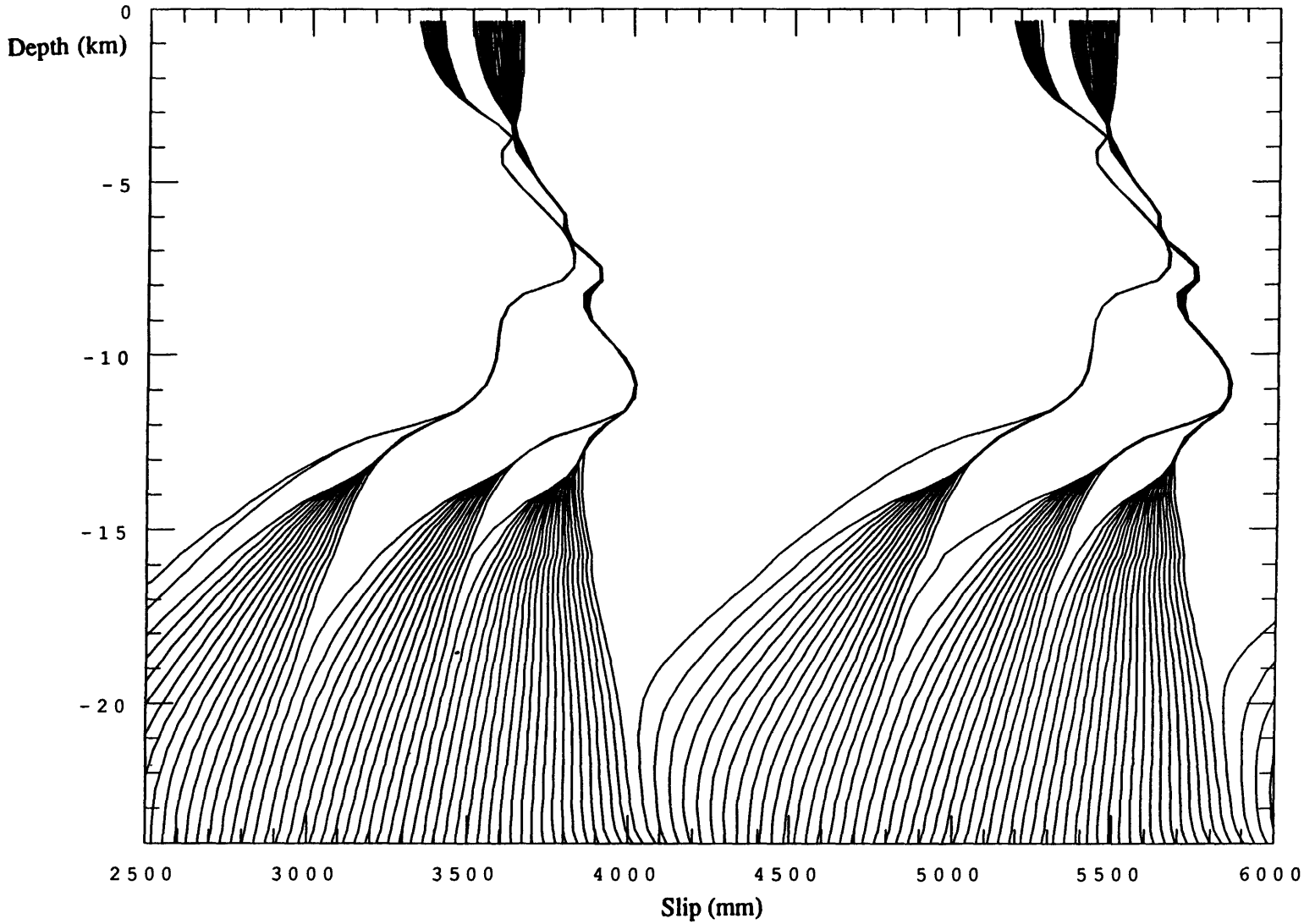



Figure 11e. Slip versus depth at 1 yr intervals for case V of text; four steps of $p$, each to 10 MPa offset from $\sigma_n$, at 4, 8, 12 and 16 km (Figure 10d).

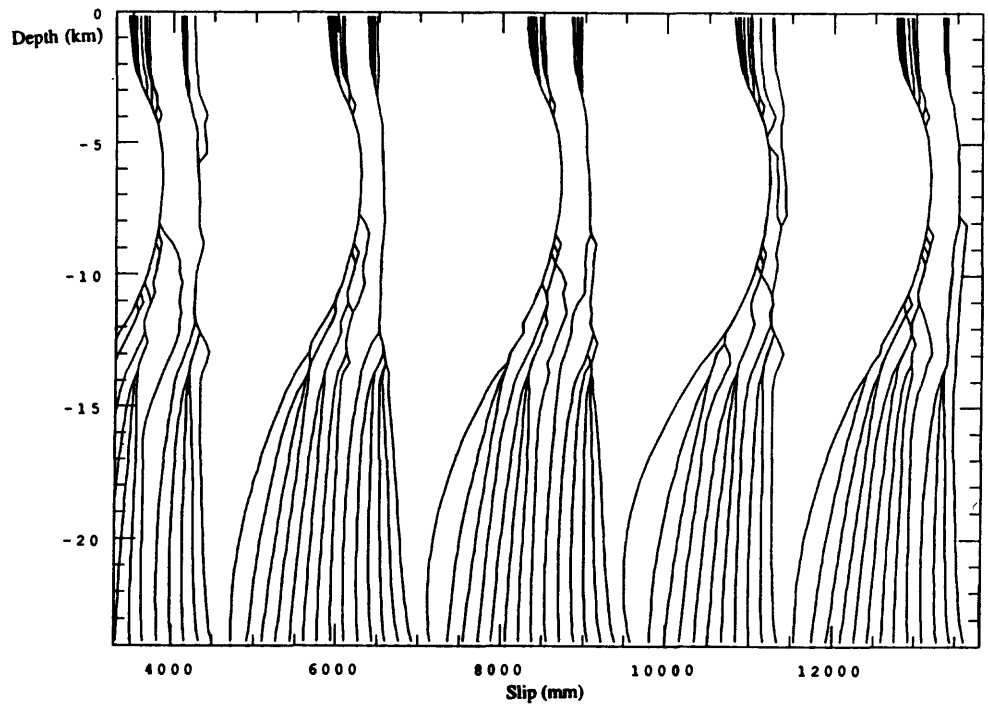



Figure 12a. Slip versus depth at 5 yr intervals for case V of text; oversized cells ($h^* = h/4$ in velocity weakening region); $p$ has lithostatic gradient at depth and is offset 50 MPa from $\sigma_n$ (Figure 10b).

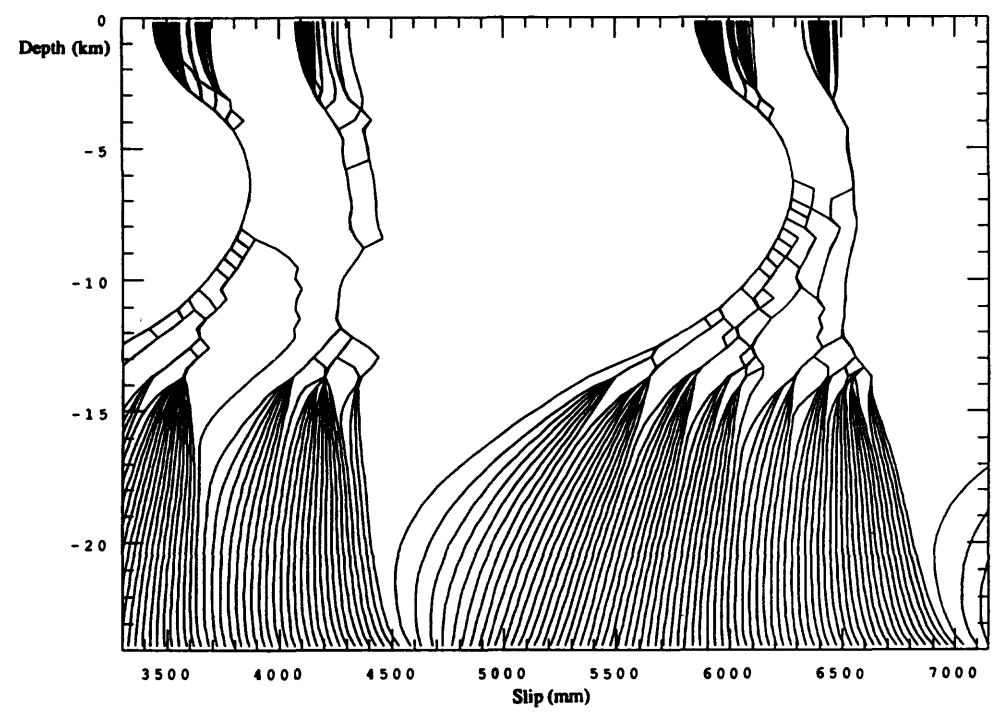



Figure 12b. Same as (a) at 1 yr intervals.

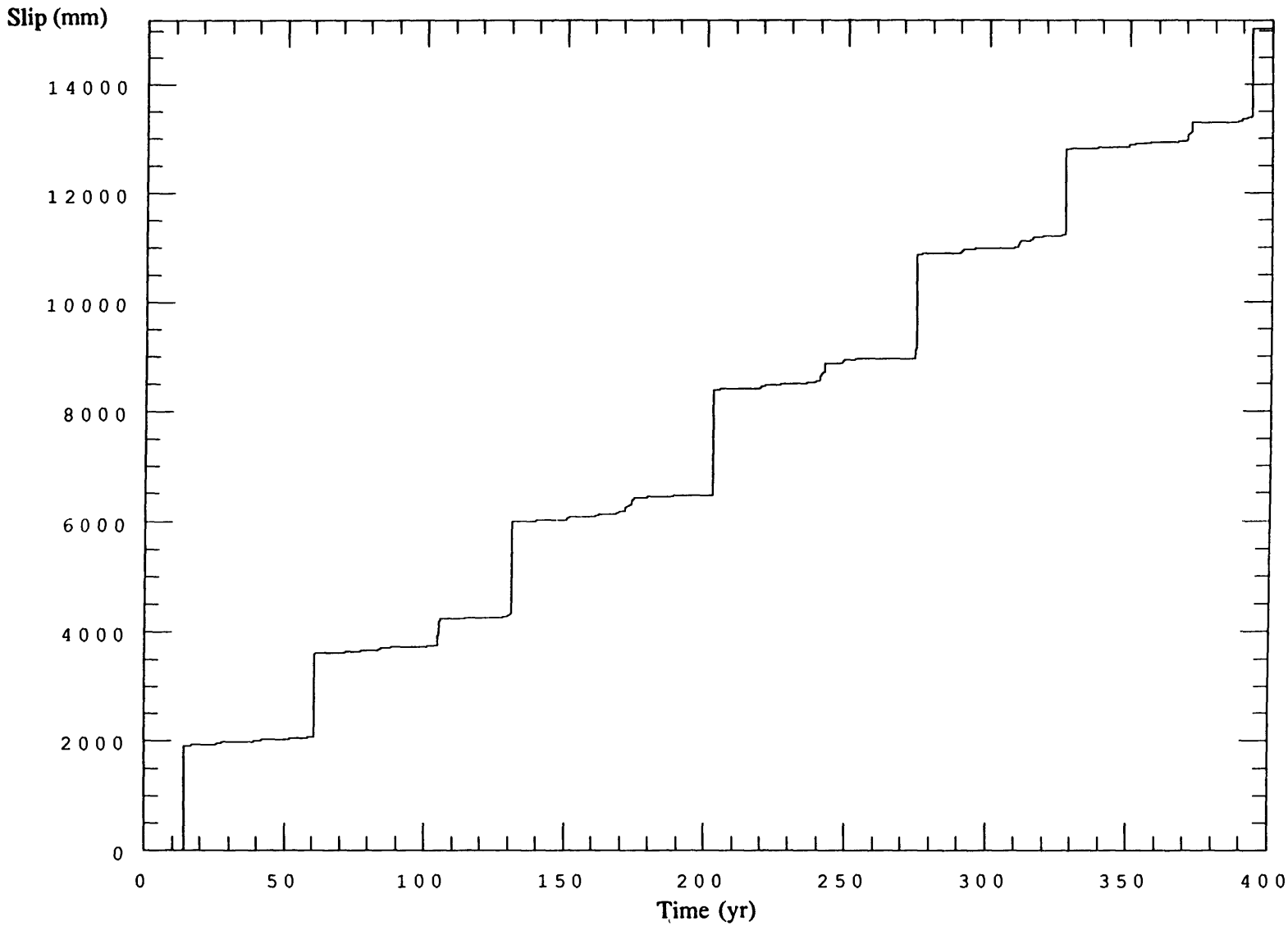



Figure 13. Average slip over the depth range 4-13 km versus time for case VI of text.

# A Friction-Feedback Model for Recent Earthquakes at Parkfield, California

**P. E. Malin [1], V. G. Oancea [2], E. Shalev [1], and T. Laursen[2]**

**[1] Department of Geology,**
**[2] Department of Civil & Environmental Engineering,**
**Duke University,**
**Durham, NC**
**27706**

Since 1990, microearthquake activity at Parkfield, California, has generally increased, with numerous events taking place on previous quiet patches, or asperities, of the San Andreas fault. Parts of this activity are organized into migrating earthquake sequences with propagation speeds much slower than possible in fault interactions based on simple elasticity and frictional sliding. Instead, the sequences can be modeled by a coupled, 2-layer, block-slider system governed by interacting velocity-weakening and velocity-strengthening frictional forces. The coupling between the two layers is assumed to be viscoelastic, accounting for aseismic slip on ductile portions of the fault. To represent the Parkfield asperity, a set of blocks in the velocity-weakening layer were given greater frictional resistance. This friction-feedback model shows how the asperity forms a barrier to the migrating sequences and how one aspect of plate motion redistributes its slip along this part of the San Andreas fault.

The San Andreas fault at Parkfield, California, continues to be an important focus of the U.S. program in earthquake prediction (1; Figure 1a.). Over the past year or so this area has experienced a notable increase in earthquake activity, including an m=4.7 event that triggered a public earthquake warning (2). This activity has begun to fill in previously quiet portions of the San Andreas fault from Middle Mountain southward down to the general region of Gold Hill (Figure 1). As in our previous study of 1988 to 1991 Parkfield earthquakes (3), some of the current activity appears to be organized in time and space, with rates and distances of earthquake migration that cannot be explained in terms of redistribution of fault slip by elastic loading of asperities held together by simple friction. Instead, the current seismicity suggests an interactive process of seismic and aseismic slip within and below the seismogenic portion of the fault.

In the year and half prior to late 1990, the Parkfield area experienced a general decline in seismicity, with several zones totally lacking in seismicity (4; Figures 1 and 2). Since 1990, the seismicity has increased, as can be seen by comparing Figures 1b and 1c. These same figures also show that portions of previously aseismic zones have now filled in with numerous small earthquakes, with considerable activity at depth below Middle Mountain and to the south near Gold Hill.

In 1990 we observed that, following a 4- to 8-month period of particularly low activity, a roughly 12-month long increase in seismicity moved more than 20 km southward along the San Andreas fault (3). The propagation speed of this migrating activity was a few tens of kms per year. After the passage of this earthquake front, the seismicity returned to former levels. The migrating front can be seen in (a) regionalized, relative seismic moment versus time, as shown in Figure 2; and (b) plots of event latitude and size versus origin time, as shown in Figure 3a. In each figure the 1990 migration front is indicated by the pair of arrows labeled as "2".

The data in Figure 3a also include more local seismicity sequences, with migration speeds much greater than a few tens of kilometers per year. In the latter half of 1990, for example,

these sequences extend 4 kms or more along the San Andreas fault. Based on their moments and corner frequencies, the largest earthquakes in the 1990 sequences had rupture surfaces less than a few hundred meters in diameter, much less than the spatial extent of the whole sequence (5). In the same time period no large earthquakes took place anywhere in the Parkfield region. Thus it appears that neither the more rapid nor slower sequences of 1990 are part of a typical earthquake aftershock or foreshock series. It is also notable that both types of sequences occurred north of the Parkfield earthquake nucleation zone near Middle Mountain (2).

We propose that the migrating earthquake sequences at Parkfield are due to the interaction of brittle and ductile failure, in large part near the base of the mostly brittle, or seismogenic, zone (6-8). The model we propose assumes a weak San Andreas fault that is strongest near this base, below which the fault is mostly ductile and weakens relatively slowly with further depth, as illustrated in Figure 4a. Both the brittle and ductile zones are subject to the strike-slip forces of plate motion, with failure of both zones tending to occur near their boundary, where this system is the strongest. At Parkfield this picture is supported by the fact that the largest microearthquakes with the largest stress drops tend to occur at this depth (4, 9)

To show how this model accounts for the migrating earthquake sequences, we have reduced it to a system of simple block-sliders, shown schematically in Figures 4b and 4c (10-12). The model consists of two, coupled, one-dimensional layers of spring-connected block-sliders with different frictional properties. In this representation, after being loaded to failure, the upper layer of blocks resists sliding with a velocity-weaking friction law. In the lower layer, the frictional resistance is assumed to be increasing with the velocity of the blocks. The blocks of each layer are coupled together by a system of dashpot-spring units (Maxwell bodies). The dashpot-spring units represent the viscoelastic character of the two crustal layers, accounting not only for their elastic connection, but their relaxation properties as well.

The specific system of blocks for our Parkfield model consists of 100 blocks: 50 brittle and 50 ductile, numbered from south to north. One end of the 2-layer system was assumed to represent the creeping San Andreas fault north of Middle Mountain. The steady pull of the lithospheric plates was applied to the ductile layer at this end. The southern end of the system was assumed to extend south of the Gold Hill area. The static friction of the entire system was randomly varied by a maximum of 10%. To account for the presumably stronger, aseismic Parkfield "asperity" near Middle Mountain where the previous M=6 earthquakes nucleated (13), the static friction of brittle blocks 15 to 22 was increased by an average of 60%.

The equations of motion in nondimensional form were derived from Lagrange's equation. These were numerically integrated for many thousands of time steps and hundreds of cases of relative strength, stiffness, relaxation time, and velocity-weakening and -hardening rates (12). Failures and rapid slip of the blocks in the brittle layer were identified as microearthquake events, whose total area and slip were used to calculate relative magnitudes. The location and size of the events were then displayed in the same fashion as the latitude-magnitude-origin time plot of the Parkfield events, as shown in Figure 3. In the case shown in Figure 3b, the dashpot-spring unit had a relaxation time of 0.5 time units and an instantaneous stiffness approximately an order of magnitude large than that of the 2 layers.

The main feature of our block-slider simulation of Parkfield seismicity is the appearance of event sequences with migration speeds below simple brittle redistribution of the plate slip. The differing speeds result from the damped feedback between the frictional properties of the brittle and ductile layers. Evidently, parts of this 2-layer, friction-feedback system can hang up, producing a period of low activity followed by a slow, southward migrating seismicity front. Rapidly moving fronts also appear, in both the north and south directions, as in recent Parkfield seismicity. The different migration speeds are a consequence of slip energy snaking back and forth between the brittle and ductile layers.

The blocks that model the Parkfield asperity form a barrier to the southward migrating sequences: Events either terminate or propagate northward from there. Presumably a simultaneous failure of the asperity blocks would represent a repeat of the Parkfield earthquake. Some activity does take place within the asperity and to its south, as slip in the ductile zone passes it by. Either the strength of the brittle zone is too high or the load too slow and small. Moreover, both the system of block-sliders and the method for calculating event size have natural limits that may not cover this case.

We recognize that seismicity at Parkfield is more complex than the activity in our simple friction-feedback model. The processes controlling the real case are likely to include vertical structure as well as chemical and rate effects not covered by our simple system. The model does suggest that recent events at Parkfield are due to the interaction of seismic and aseismic slip around the asperities of the San Andreas fault there. The occurrence of microearthquakes within and south of these strong patches reflects the continuing load on them and provides a means of observing slip changes at depth beneath them. The time and space character of these events before the next Parkfield earthquake should resolve whether this event was part of a migrating stress front or an isolated failure.

## References and Notes


1. A. Michael and J. Langbein, *EOS* **74**, 145, 153 (1993).
2. J. Langbein et al., *ibid* **74**, 152 (1993).
3. P. E. Malin and M. G. Alvarez, *Science* **256**, 1005 (1992).
4. P. E. Malin et al., *ibid* **244**, 557 (1989).
5. The moments and corner frequencies were determined from the S-wave frequency spectra using the methods of D. J. Andrews, in *Earthquake Source Mechanics,* K. Aki and P. Richards, Eds., *American Geophysical Union Monograph* ***37,*** American Geophysical Union, Washington, D. C., (1986). The largest earthquake in 1990 was a local magnitude 3.3; the m=4.7 took place on Oct. 20, 1992, rupturing a few kilometers of fault and producing a more typical aftershock sequence (2).
6. J. Savage, *J. Geophys. Res* 76, 1954 (1971).
7. F. K. Lehner, V. C. Li, J. R. Rice, *J. Geophys. Res* **86**, 6155 (1981).
8. R. Wesson, *EOS* **65**, (1985).
9. M. E. O'Neill, *BSSA* **74**, 27 (1984).
10. J. Haung and D.L. Turcotte, *Nature* **348**, 234 (1990).
11. B.E. Shaw et al., *J. Geophys. Res.,* **97,** 479 (1992).
12. Malin and Oancea, ms in preparation (1993).
13. W. H. Bakun and T. V. McEvilly, *J. Geophys. Res.* **89**, 3051 (1984).


14. We thank colleagues at the University of California at Berkeley and the U.S. Geological Survey for sharing the effort of keeping the Parkfield downhole seismology project alive. This work was support by DOI-USGS grant #14-08-0001-G1375.

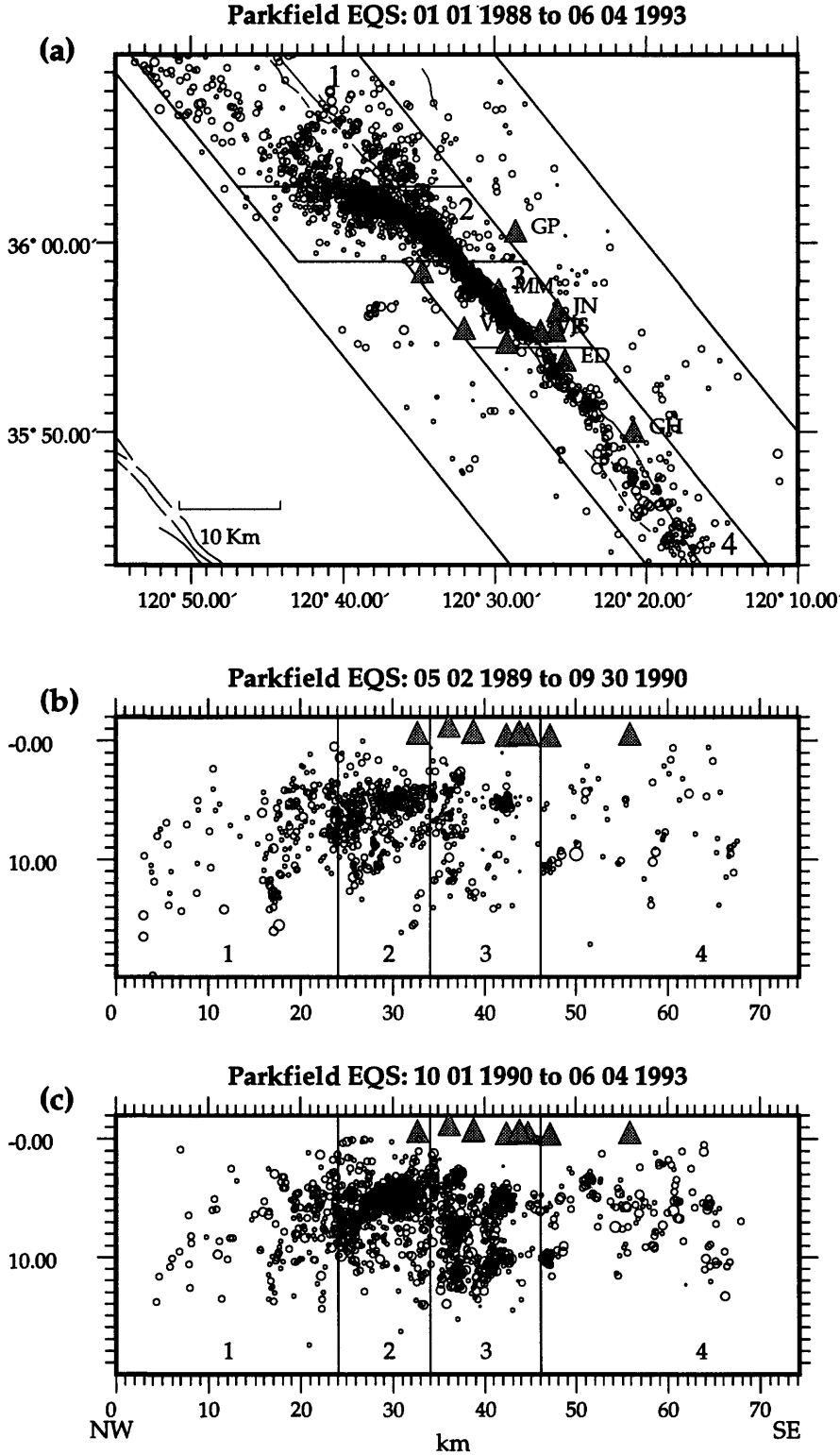



Fig. 1. Part (a). Map of the seismicity along the San Andreas fault at Parkfield, California. The downhole seismic stations used in this study are shown as triangles, the Middle Mountain station at MM and the Gold Hill station at GH. Also shown are regional boxes 1 to 4 used in calculating the regionalized seismic moments of Figure 2. Parts (b) and (c). Cross sections along the San Andreas fault at Parkfield, California, showing the regional boxes and hypocenters of microearthquake in 2 different time periods. The later time period in Part (c) shows a significant number of microearthquakes taking place in the previously quiet patches seen in Part (b).

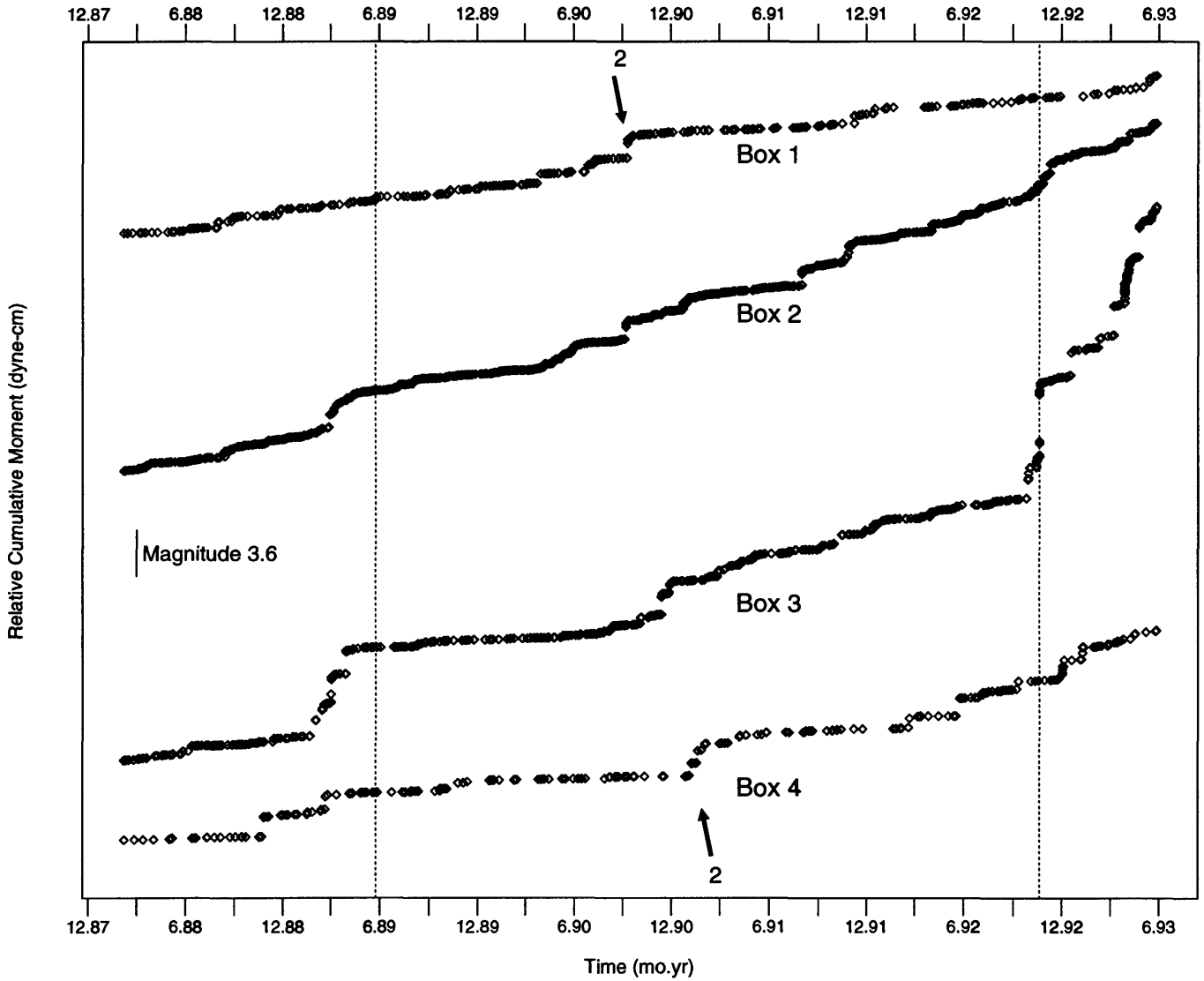



Fig. 2. The relative cumulative moments of Parkfield earthquakes by region, illustrating the migration of earthquake-related slip north and south along the San Andreas fault. The M~4.1 and M~4.7 events of May 29, 1989 and October 20, 1992 were removed from the moment contribution at the time shown by the vertical dotted line. Arrows labeled #2 indicate the slow migrating earthquake sequence of 1990. The more rapid local sequences are averaged by the size of the boxes.

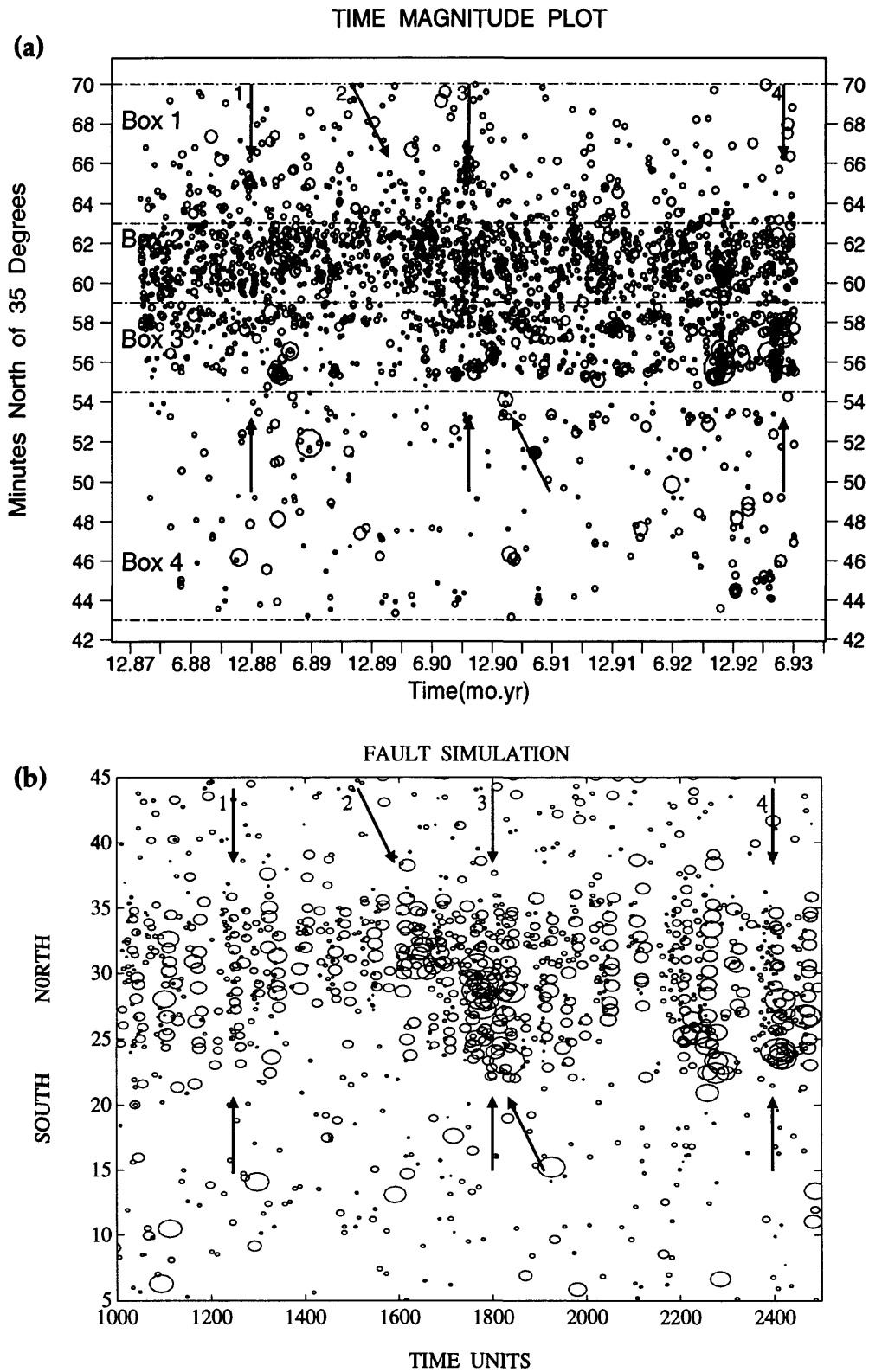



Fig. 3. Part (a). Earthquake magnitudes along the San Andreas fault as a function of latitude and time. The magnitudes of the events are represented by circles of different sizes, with the October M~4.7 earthquake being the largest and M~0 the smallest. Arrows point out the proposed slow (pair 2) and rapid earthquake migration sequences (pairs 1, 3, 4). Part (b). Results of the friction-feedback model of Parkfield seismicity. As in Part (a), the sizes of the ellipses represent the relative sizes of the events. As indicated by the arrows, both slow, regional and rapid, local earthquake sequences can be simulated with this type of model.

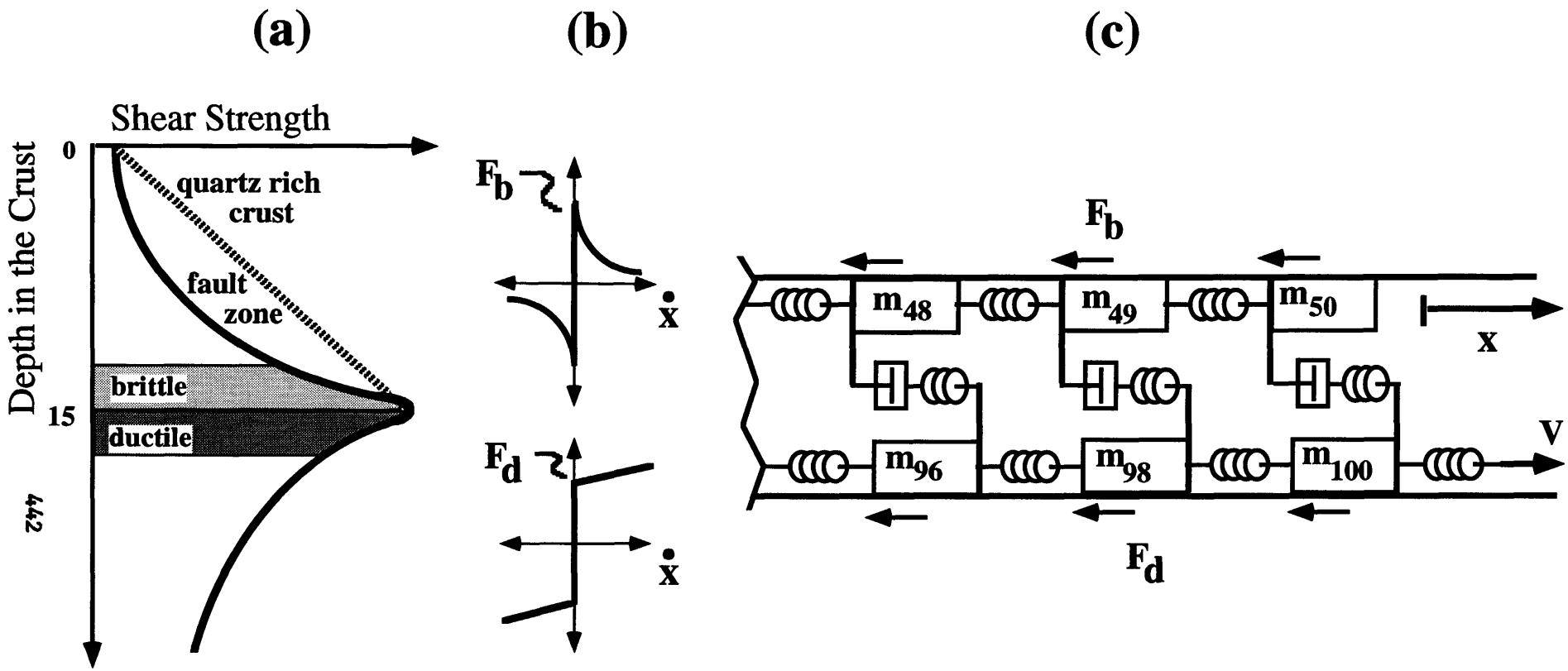



Fig. 4. Part (a). A hypothetical strength-verses-depth curve for the San Andreas fault at Parkfield. The curve is based in part on numerous events and higher stress drops near the bottom of the seismogenic zone (4, 9). Below this level the fault begins to weaken at a relatively slow rate. Part (b). The velocity-weakening and velocity-hardening friction laws assumed for the brittle (Fb) and ductile (Fd) layers of the model (10, 12). Part (c). A schematic representation of the brittle and ductile layers at Parkfield in terms of a block-slider system of springs, dashpots, and masses.

# DYNAMICS AND FRICTIONAL HEAT GENERATION IN A FOAM RUBBER MODEL OF EARTHQUAKE STICK-SLIP

Abdolrasool Anooshehpoor and James N. Brune

Seismological Laboratory, 168
University of Nevada-Reno
Reno, Nevada 89557-0141

## ABSTRACT

Spontaneous stick-slip along the interface between stressed foam rubber blocks is a simple analog of earthquake rupture and stick-slip. There are two dynamic scaling considerations not satisfied by ordinary rock mechanics experiments which are satisfied by the foam rubber model. First, the stressing apparatus for the foam model is effectively infinitely rigid compared to the rigidity of the model. Second, the dimension of the foam model is large compared to the dimension of dynamic slip pulse (approximately 200cm vs. 10cm) consistent with the abrupt-locking slip pulse model of Brune (1970, 1976), and Heaton (1990).

Results from a study of stick-slip particle motion at the interface, indicate that normal vibrations and interface separation are an important part of the stick-slip process. In most cases, there is a clear interface separation during stick-slip. A comparison of frictional heat generation in the stable-sliding and stick-slip foam rubber models indicates a linear relation between the temperature increase on the fault surface (for a given distance of slip) and the driving shear force for the stable-sliding model, while for the stick-slip model, there is essentially no variation in frictional heat generation with increase in shear stress. The phenomenon of interface separation, if it applies to earthquakes, may thus explain some of the most frustrating problems in earthquake mechanics, *e.g.*, the heat flow paradox, the long term weakness of major active faults, and anomalous P-wave radiation.

## INTRODUCTION

In spite of years of study and speculation, the heat flow paradox remains a paradox. New dimensions of the problem have been highlighted by recent results from the Cajon Pass drill hole, and recent evidence that absolute shear stresses along the San Andreas Fault may be low. Two alternatives have been suggested in recent years: (1) high fluid pressures may exist on the fault and (2) dynamic inertial effects, including variations in normal stress

(and possible fault opening) may mean that the mechanism of stick-slip in rock mechanics experiments does not correspond adequately to the mechanism of stick-slip in earth (because of improper scaling, and lack of proper consideration of dynamics effects).

### High Pore Pressure

At present the most frequently discussed mechanism for explaining the heat flow paradox is some type of reduction of effective normal stress by near lithostatic pore pressures, either caused by mantle sources of pore pressure along the fault (Rice, 1992), or by pods of high pore pressure generated and locked in by permeability valves (Byerlee, 1990; 1992 and Rice 1992). The high pore pressure explanation for earthquake mechanism has been around in various forms since Hubbert and Rubey (1959) suggested it as the likely explanation of the paradox of large overthrusts. The law of "effective normal stress" (effective normal stress equals the actual normal stress minus the pore pressure) has been verified in numerous experiments in the laboratory. Therefore if high pore pressure exists on faults for which the stresses are demonstrated to be low, then it is likely to be the explanation for the heat flow paradox and the paradox of low shear stresses. There is evidence of local high pore pressure in many geologic environments. However, there is little direct evidence of high pore pressure on the San Andreas fault, nor on other major faults. The Cajon Pass Drill Hole Experiment did not uncover evidence of high pore pressure to depths of about 3 km, although the hole is not directly on the fault and may not have been deep enough.

A major problem with the high pore pressure explanation is the question of how the pore pressure can be maintained at such high levels for hundreds, up to tens of thousands of years (the repeat times of major earthquakes on some faults). Faults are zones of fractured rock, and consequently would ordinarily be expected to be zones of high permeability and consequent low pressure. However, various mechanisms could lead to lower permeabilities (perhaps anisotropic) in fault zones (*e.g.*, chemical solution and exsolution, and sealing by clay minerals). If pore pressure changes are to explain features of earthquakes such as rupture lengths of thousands of kilometers, rupture complexity, slip partitioning, triggering by transient static and dynamic strains, very long repeat times of up to 100,000 years, and complexities of aftershock slip orientations, many at present unknown or poorly understood parameters of fluid flow will have to be introduced.

### Dynamic Mechanisms of Stick-slip

Dynamic effects in the process of stick-slip have the potential of not only explaining the heat flow paradox, but also explaining other puzzling features of earthquake slip, some of which are important to earthquake hazard research. One explanation of the heat flow paradox that has recently been suggested is that the actual mechanism of stick-slip observed on small rock samples in the laboratory (which originally led to the heat flow paradox, as a consequence of high stresses determined to be necessary for stick-slip), may not correspond to the mechanism of stick-slip in the earth, in part because of dynamic scaling problems, sample-machine

interaction, and the existence of large variations in normal stress, and possibly even fault opening during the stick-slip process (Brune et al., 1989, 1990, 1992).

In laboratory experiments inertial forces related to the volume of rock involved are small compared to the forces acting on the fault zone. For earthquakes, however, since the inertial forces scale as the volume (dimension cubed) and fault surface forces scale as the fault area (dimension squared), inertial forces could become comparable to fault surface forces. From this point of view, if inertial effects explain the heat flow paradox, then the heat flow paradox may have arisen because of neglect of proper scaling from small laboratory samples to the dimensions involved in earthquakes. For example the inertial forces involved in fault ruptures which pass around bends or jogs of a given angle are quite small in laboratory size samples but quite large on the scale of earthquakes.

The dynamic normal vibrations mechanism was suggested by modeling of stick-slip between large blocks of foam rubber (polyurethane). In these models, opening of the fault during stick-slip is clearly observed, resulting in a consequent reduction of frictional heat generation (Anooshehpoor and Brune, 1992). Also the "rise time" or duration of fast slip in the foam rubber model is short compared to the fault dimension divided by the rupture velocity, in agreement with the abrupt-locking slip pulse models of Brune (1970, 1976) and Heaton (1990), for which there is considerable support from spectral analysis (Smith et al., 1990), and from fault slip inversion studies of recent earthquakes (Heaton, 1990).

Normal interface vibrations and interface separation may provide a solution to the heat flow paradox, i.e. the lack of any indication of frictional heat generation along the San Andreas fault (Brune, et al., 1969, Lachenbruch and Sass, 1973). If waves of separation occur during earthquake faulting, or if the normal stress is reduced, the frictional heat generation will be reduced. Anooshehpoor and Brune (1992) indicated a reduction of frictional heat generation of a factor of 5 (over the frictional heat generated when sheets of the same material were forced to slide the same distance, but kept in continuous contact).

## Partial Stress Drop—Abrupt Locking (Self-Healing) Slip Pulse Model

The length of the slip pulse can be estimated by multiplying the rise time of the shear slip pulse by the rupture velocity. The rise times for the foam rubber model (indicated in Figure 2) are approximately 2.5 ms, and the rupture velocity has been determined to be about 3000 cm/s. This gives a slip pulse length of about 7 cm. This is considerably less than the dimensions of the fault surface (about 200 cm). Thus the portion of the fault that is slipping at any given time is a small fraction of the total surface, and the rupture corresponds to the abrupt-locking, self-healing slip pulse of Brune (1970, 1976), and Heaton (1990). Brune (1970) suggested that in reality it was unlikely that a fault could slip to 100% of the dynamic stress drop over the complete fault plane, and that the consequent "partial stress drop" model would produce an intermediate spectral slope of approximately $\omega^{-1}$. It was suggested that this could be caused by a complex, multiple event stress drop, or a type

of slip from one "potential well" to another as would be appropriate for a crystal dislocation where a molecule on one side of the dislocation slips over another on the other side. This mechanism might be appropriate for interlocking asperities on a fault. Brune (1976) called this the "abrupt locking"model. Papageorgiou and Aki (1983), and Papageorgiou (1988) considered a special case of this type where uniform circular patches were allowed to slip, with the fault pinned by "barriers" between circles. For a review of the evidence for the partial stress drop model see Brune et al., (1990).

In the time domain, the partial stress drop model has the stress drop and slip velocities temporarily and/or locally higher than would be the case if the final static drop had been applied permanently and uniformly. Thus the average stress must first drop and then increase, but not back to the original stress level, leaving a permanent average stress drop smaller than the transient stress drop. This model for earthquake stress was suggested by Housner (1955). The high initial (or local) stress drop (difference between the initial stress and the sliding friction) is often referred to as the dynamic stress drop (as compared to the usually lower static stress drop). Heaton (1990) calls this model the self-healing slip pulse model (the fault heals and leaves a final stress level higher than the sliding frictional stress). Quin (1990) refers to this as a "moving window of radiation".

Heaton (1990) has documented that most of the recent earthquakes which have allowed detailed determination of dislocation time histories by inversion techniques have shown local rise times for fault slip much shorter than would be the case for the uniform stress drop model (in which slip near the center of the fault continues until a"healing" signal arrives from the edges, giving a relatively long rise time). The data he presented indicate that a complex multiple event or multiple asperity stress drop model is appropriate for nearly all these events, and that short rise times are the rule rather than the exception.

### Anomalous P-wave Radiation

If there are tensile stress or tensile opening motions on the fault, this may lead to anomalous P-wave radiation. Haskell (1964) suggested that tensile-like normal interface vibrations on the fault plane were required because the radiation of P-wave energy from large earthquakes was too high for pure shear faulting. Monlar et al., (1973) found higher corner frequencies for P waves than S waves, as well as anomalously high P-wave energy beyond the corner frequency. These results are unexpected for pure shear faulting at near S-wave rupture velocities (Dahlen, 1974). Blandford (1975) proposed a spectral model containing of both shear sources and tensile sources, and was able to increase the ratio of compressional wave to shear wave energy. Evidence from the ANZA array, (Vernon et al., 1989) and Guerrero digital strong motion array (Castro et al., 1991) gives some indication of anomalous P-wave excitation which might be produced by normal interface vibrations. Unfortunately the data are subject to large uncertainties because of unknown effects of attenuation (which might affect the S-waves differently than the P-waves). Walter and Brune (1993) presented an average model for the far-field spectrum of tensile cracks and superpositions of shear cracks

and tensile cracks. Brune et al.,(1992) provided preliminary evidence of high frequency P-wave radiation from the fault surface in the foam rubber model.

### Schallamach Waves

The most familiar interface wave involving separation are Schallamach waves which occur when two media of large difference in rigidity slide past one another (Schallamach, 1971), for example when tires squeal against pavement. The physical principle involved is that it is more energy efficient to create a wrinkle in the softer material, allowing separation along the interface, and thus sliding without frictional contact. Barquins and Courtel (1975) concluded that Schallamach waves of detachment are governed by relaxation effects and adhesion between the two materials. Schallamach waves can be quasi-static or dynamic, depending on the velocity of sliding.

### Comninou-Dundurs Waves

Comninou and Dundurs, in a series of theoretical papers (1977, 1978a, 1978b) proposed that when two media with the same properties are pressed together and sheared at the same time, a dynamic interface wave involving separation could be a stable solution to the equations of motion and boundary conditions. These waves propagate faster than Schallamach waves of detachment. They are controlled by elasticity and inertia and have a phase velocity which may have any value falling between the velocities of Rayleigh and shear waves. The validity of the solution was questioned by Freund (1978), and the solution has been ignored by seismologists and rock mechanics researchers. The aspect ratio of the opening of these waves decreases with increasing normal stress. It appears that Freund's (1978) objection is valid, and thus Comninou-Dundurs waves probably cannot propagate without some mechanism of transferring energy from the back of the crack to the front. However, since opening waves are observed in the foam rubber model, some modification of Comninou-Dundurs theory incorporating transfer of shear momentum into tensile momentum may be possible. Zeng et al., (1992) calculated interface separation caused by asperity impact, and concluded that interface separation could be caused by local asperity impacts with stress concentrations of a factor of 100.

### Nucleation

If dynamic inertial forces can explain the heat flow paradox and the dynamic weakness of faults, the problem of nucleation of rupture remains unexplained. In this context we may define nucleation as the process which causes slip to accelerate from low velocity slip which is quasi-static (inertial forces can be ignored) to dynamic (inertial forces can weaken the fault). There are two ways in which this transition might occur. The first is to concentrate stresses to near the failing point of rock (kilobars) over a local zone sufficiently large so that when the zone fails it releases enough kinetic energy to initiate a dynamic rupture which will continue to propagate because of dynamic effects. The second is to have slip gradually

accelerate (due to some fault weakening mechanism) until the velocities are sufficiently high to make the transition from the quasi-static to dynamic regime.

Several mechanisms might lead to high stress concentrations over fairly large zones, *e.g.*, progressive weakening of certain parts of the fault might concentrate stress on stronger parts. Progressive weakening might be brought about by changes in pore pressure on portions of the fault, by chemical solution and exsolution, or by permeability sealing and opening. Once stress is concentrated to a high level over a sufficiently large zone, subsequent failure could release enough energy to result in the transition from the quasi-static to dynamic regime.

Three mechanisms leading to pore pressure changes which might cause rupture nucleation have recently been suggested. The first is pore pressure gradients between the fault zone and surrounding medium caused by fluid pressures generated in the underlying mantle (Rice, 1992). The second is development of compartmentalized zones of lower and higher fluid pressures caused by permeability gradients and pressure gradients developed quasi-statically, and locked in by permeability valves which can break, leading to relatively rapid quasi-static concentrations of stress and consequent dynamic failure (Byerlee, 1993). The third mechanism is creation of high pore pressure by compaction of fault gouge (Sleep and Blanpied, 1992). Fault weakening during the faulting process might be caused by complex fluid flow effects, by local melting, or by creation of steam pressure from fault heating (Lachenbruch, 1980).

## DESCRIPTION OF FOAM RUBBER MODEL AND RESULTS

The model consists of two large blocks of foam rubber ($\sim 8 \times 9 \times 3 ft^3$), one driven horizontally over the other by a hydraulic piston (Figure 1). At the contact (fault), the blocks are locked by gravity acting as a normal force (the normal force is increased or reduced by screws attached to the upper surface. As the upper block is forced over the lower block, the strain in the blocks increases until the stress at the interface exceeds the frictional resistance and a stick-slip event occurs over the whole boundary (fault plane). These major events correspond to "characteristic" events for the system, analogous to "characteristic earthquakes" in the earth. Successive "characteristic" events usually cause about the same amount of average slip between the blocks, but the pattern of slip can vary markedly, with the rupture initiating at different points and propagating in different directions.

If the driving displacement is steady, the characteristic events repeat more or less regularly until the upper block has slipped about 50 centimeters, corresponding to about 50 characteristic events and some additional smaller events. At this point, the hydraulic piston is fully extended and one experimental run is complete. The stress is relieved and the upper block moved back to the starting position for repeat of the procedure.

The displacement as a function of time of the upper block can be monitored by a pen (attached to the upper block) moving over a strip chart recorder, or by electronic position

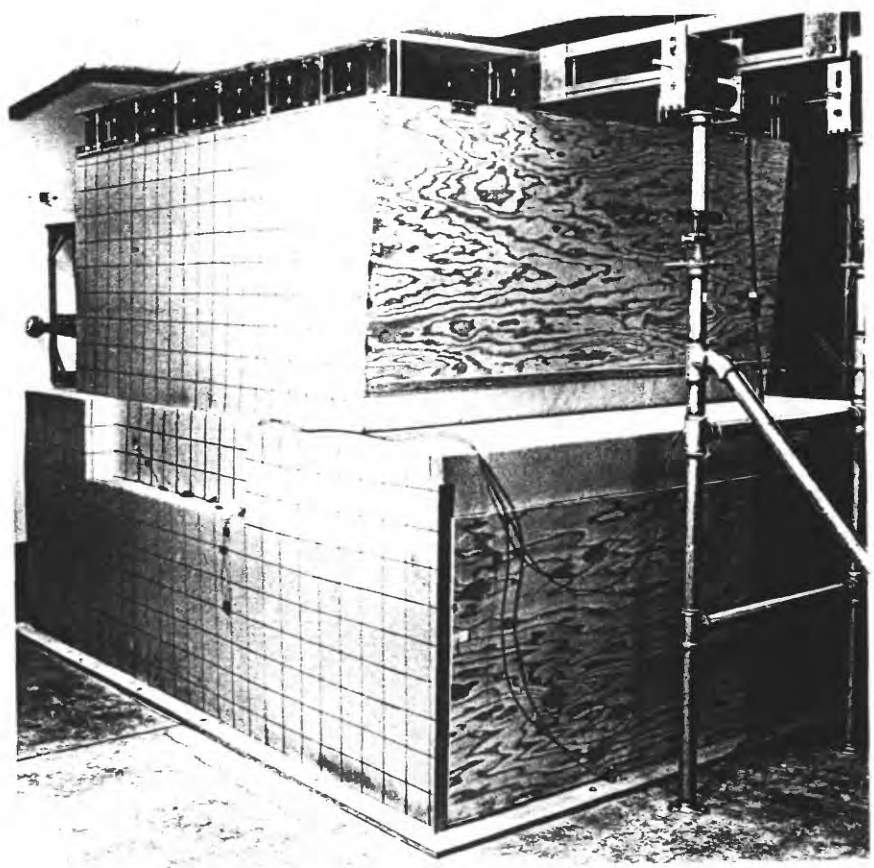


Figure 1: The model consists of two large blocks of foam rubber, one driven horizontally over the other by a Hydraulic piston. Normal force at the interface is provided by gravity, and can be controlled by screws holding up the upper block

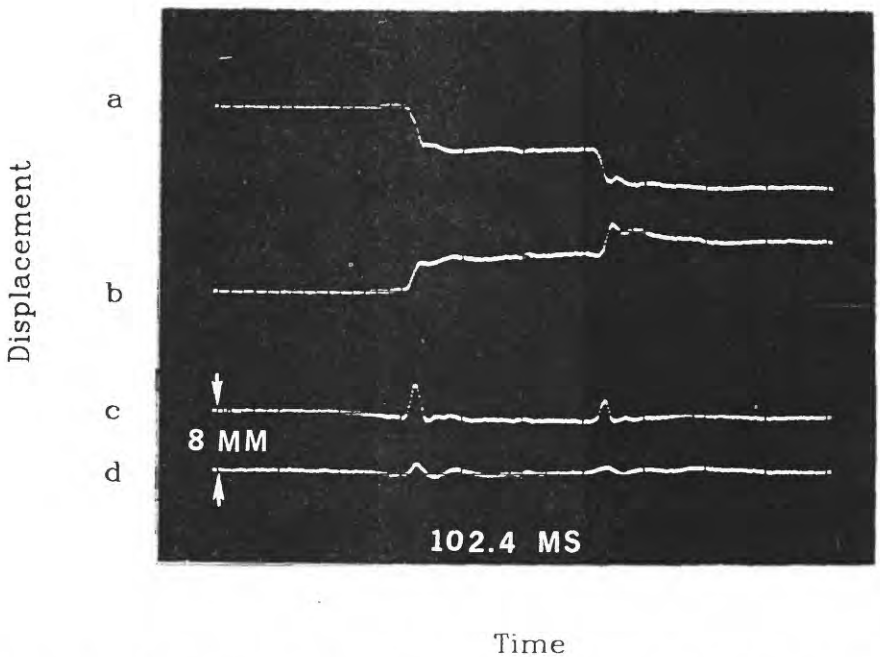



Figure 2: Shear (a & b) and normal (c & d) components of particle motions at two locations above and below the interface in foam rubber model of stick-slip. The normal motion of the upper block is larger than the lower one implying interface separation.

detectors. Seismic events can be monitored by accelerometers embedded in the foam near the fault surface. The waveforms are recorded both on digital Biomation transient recorders and on a PC with a data acquisition system. The absolute stress and stress drop during seismic slip is monitored by an in-line hydraulic pressure gauge or by a scale on the piston system. The stiffness of the driving system is much greater than that of the foam rubber, assuring that the driving system does not interact with and perturb the responding system.

## Short Term Predictability and Premonitory Phenomena

Most hopes for reliable short term earthquake prediction are based on some form of accelerating change in observable conditions immediately prior (hence, short term) to a major earthquake (the "characteristic earthquake"). Such accelerating changes in conditions might be related to high stress and consequent non-linear behavior, *e.g.* ductile deformation, dilatancy, etc. In simple systems with planar surfaces and essentially linear deformation, such as the granite block system of Dieterich (1986) and the foam rubber block system, the only likely premonitory phenomena are related to accelerating creep. The foam rubber model may be stressed to very near the peak value for nucleation and remain stable indefinitely if no more stress is added. Thus there are no time-dependent non-linear creep phenomena which drive the system toward instability without the addition of more stress. Once an instability develops, the rupture grows to the edges of the model. In this simple system, short term prediction has to be based on the characteristics of the nucleation of unstable patches.

Parameterization of the constitutive properties of fault slip has been the subject of a series of recent analytical and experimental studies by a number of authors (Dieterich, 1979, 1981; Ruina, 1980, 1983; Okubo and Dieterich, 1986). A recent paper by Dieterich (1986) builds on the results of these studies to present a model for the nucleation of earthquake slip with simple predictions which can be compared with our foam rubber model results.

For the foam rubber model the characteristic roughness ($D_c$) (Dieterich, 1986) is about 1 mm, the typical size of the vesicles in the polyurethane lattice, and the typical displacement required before stick-slip can be initiated. Thus by analogy we might expect premonitory slip of the order of a few millimeters. The critical patch radius in the Dieterich model is given by equation 14 in his paper (Dieterich, 1986). Assuming typical strain drops of about 10% we infer that the critical patch radius for instability is about 10 cm, consistent with results for our model. Most if not all stick-slip events we observe occur over surfaces larger than 10 cm and have displacements of at least a few millimeters. A further check on the analogy with the Dieterich granite block results is provided by direct observation of fault motion using velocity sensors embedded in the foam rubber a few millimeters away from the fault surface, so that they are primarily sensitive to fault slip next to the sensor. We have observed numerous stick-slip events(see Brune et al., 1992). For most of our observations the nucleation zone and associated premonitory slip was too small to be captured by our limited array of sensors, consistent with the estimations above of critical nucleation patch size of about 10 cm. Thus we conclude that the foam rubber model behaves qualitatively like the granite blocks of Dieterich (1986).

## Scaling Considerations in Understanding Stick-slip

There are two dynamic scaling considerations not satisfied by ordinary rock mechanics experiments which are satisfied by the foam rubber model. First, the stressing apparatus for the foam model is effectively infinitely rigid compared to the rigidity of the model. Second, the dimension of the foam model is large compared to the dimension of dynamic slip pulse.

*Rigidity of Stressing Apparatus*— The stiffness of the stressing apparatus in the foam rubber model is given by the compressive deformation modulus of the steel and hydraulic fluid. We have verified that this is effectively infinite compared to the stiffness of the foam rubber by placing a small accelerometer on the piston and verifying that at the time of a stick-slip event there is no motion of the piston.

*Dimension of the Slip Pulse*—As described above the overall dimensions of the foam rubber model are large compared to the dimensions of the dynamic slip pulse which propagates along the interface between the two blocks. This allows the slip pulse to propagate predominantly under the influence of conditions local to the slip pulse itself, with minimized effects of the boundaries of the model and the stressing apparatus. This obviously corresponds better to the conditions in the earth, for which the length of slip pulse is small compared to the

dimensions of the fault (Heaton, 1990). In rock mechanics experiments this is not the case. Even in the case of the large granite blocks of Dieterich (1986), the rise time multiplied by the rupture velocity, $\sim$ 5 meters, is considerably larger than the dimensions of the model. The situation is even more extreme in the case of small size rock mechanics samples. Thus the effects of the stressing apparatus or the boundaries of the sample are necessarily directly involved in the dynamics, and this makes the analogy with actual earthquakes questionable. In this respect the foam rubber model corresponds better to the real earth than do small rock models of stick-slip in the laboratory.

### New Stable Sliding Foam Rubber Model

We have recently discovered that stable sliding (without stick-slip) occurs in a foam rubber model with properties only slightly different from those which generate consistent characteristic stick slip events. This was discovered because we purchased two blocks of foam rubber from a different company which used a different foaming process, involving freon. The surface feels slightly different to the touch, has a slightly higher rigidity, and probably slightly different interface properties. This model exhibit many small stick slip events when the normal stress is near zero, but once the normal stress is increased slightly, stick-slip does not occur, only stable sliding. Comparison of results from the two types of models should eventually enable us to elucidate some of the factors controlling the stick-slip and stable sliding regimes.

Previous studies of the transition from stable sliding to stick-slip in laboratory models have shown that in some cases very small changes in interface conditions may change behavior from stable sliding to stick-slip. Stephen Brown (personal communication) found that in a polycarbonate model of stick slip (Brown et al., 1991) a slight brushing of the surface with emery cloth was enough to transform the behavior from stable sliding to stick-slip. In the earth it is believed that at shallow depths, particularly in soft sedimentary rocks, faults stably slide because the normal stress is low and the surface energy density is not high enough to lead to stick-slip. With increasing depth and normal stress the fault becomes "seismogenic", *i.e.*, the fault moves from a stable sliding regime to a stick-slip regime, similar to transition observed in many rock mechanics experiments. At depths below the seismogenic zone it is believed that temperature and pressure reach such high values that rock creep rather than brittle behavior takes over, and stable sliding occurs, again similar to processes observed in the laboratory. The new stable sliding foam rubber model may be analogous to the latter case, since the measurements indicate a decreasing co-efficient of friction with increasing normal stress. At any rate, comparisons of frictional heat generation in the stable sliding model with that in the stick-slip model will elucidate the role of stick-slip in reducing the strength of the surface and in reducing the frictional heat generation.

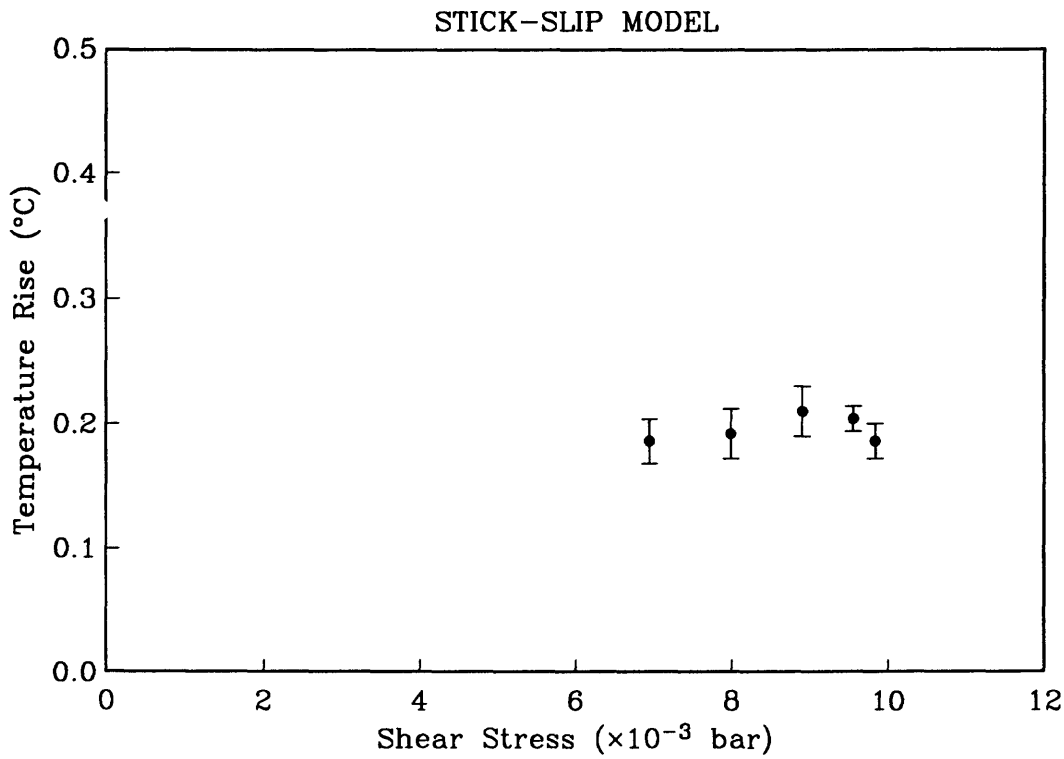



Figure 3: Temperature on the fault is measured by placing thermistors near the fault surface. The change in temperature for a single event is insignificant. In order to have a measurable change, the upper block is moved over the lower block for a total distance of 45 cm. The change in temperature on the fault is virtually independent of the normal stress for the stick-slip model.

## Comparison of Frictional Heat Generation in the Stable-Sliding and Stick-Slip Foam Rubber Models

We have measured the frictional heat generation in both the stable sliding and stick-slip foam rubber models by inserting small thermistors near the interface between the two blocks of foam rubber (Anooshehpoor and Brune, 1992). The heat generated by the stick-slip model is considerably less than that generated when two thin sheets of the same material are forced to translate the same distance in frictional contact. Thus it appears that the stick-slip process reduces the frictional heat generation. Anooshehpoor and Brune(1992) suggested that this is a direct result of the interface separation involved in the stick-slip process in foam rubber. On the other hand, for the stable sliding model no seismic energy is radiated, and thus the frictional heat generation must reflect the total amount of work done on the model.

In both types of models we have made measurements of frictional heat generation for different normal stresses. Figure 3 shows the temperature rise on the fault surface (for a total slip distance of 45 cm) for the stick-slip model. There is essentially no variation in frictional

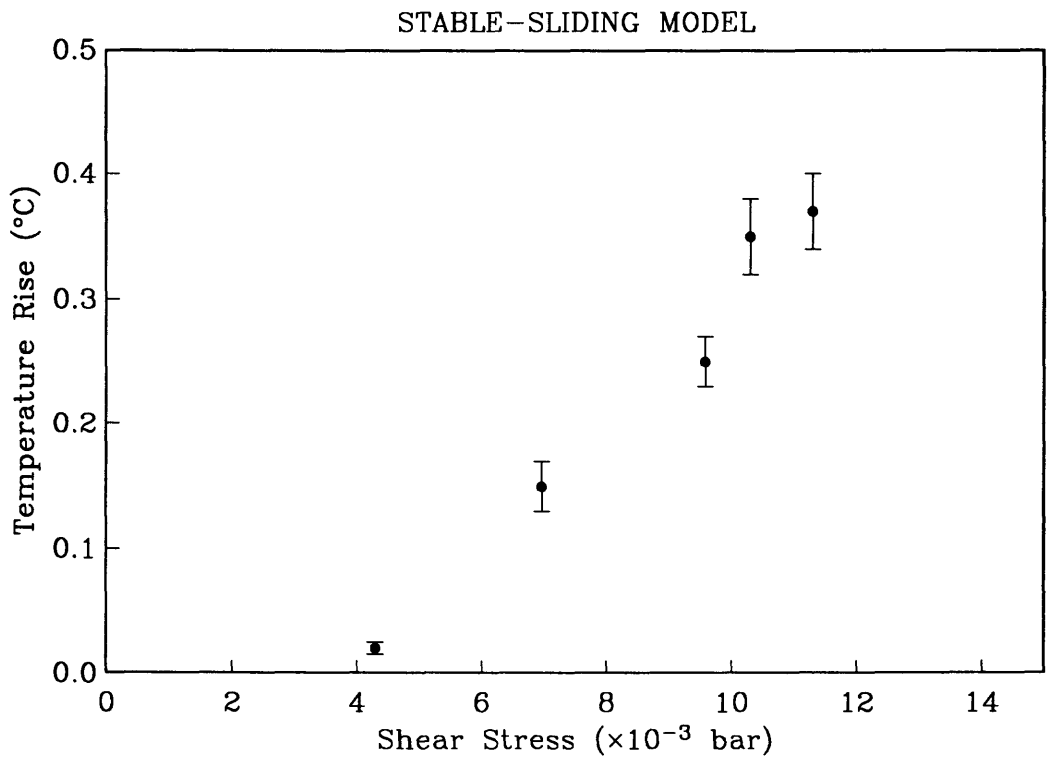



Figure 4: For the stable-sliding model, the temperature change on the fault for a given distance of slip is linearly proportional to the shear force.

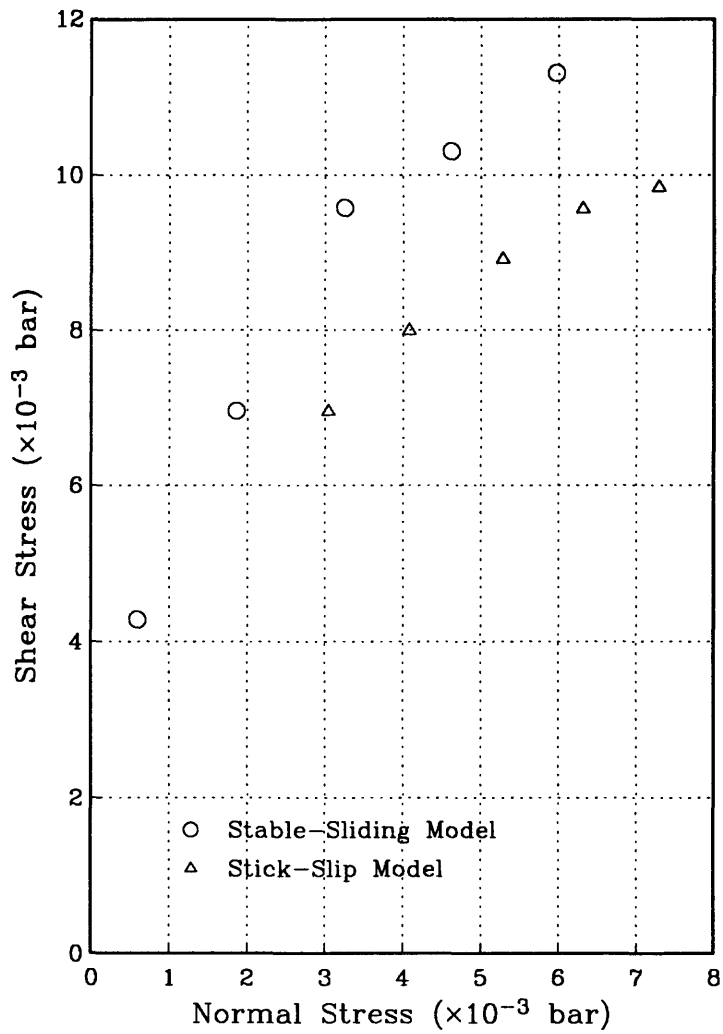



Figure 5: A plot of shear stress vs. applied normal stress in both models. In both models,the effective coefficient of friction decreases with an increase in normal stress.

heat generation with increase in normal stress, even though the shear stressing force increases (Figure 5) with normal stress. This would seem to be a very odd result in light of commonly assumed models of stick-slip, implying that the effective coefficient of friction varies inversely with the normal stress. However, in the light of interface separation during fault slip this result is clearly understandable, since most of the stick-slip occurs when the interface is open. Preliminary results show that by increasing the normal stress at the fault, the normal (to the fault surface) component of motion increases. This suggests that the increase in radiated seismic energy with increase in normal stress occurs mostly as P-wave radiation.

For the stable sliding model there are no dynamic effects to reduce the frictional heat generation, and thus we expect the heat generation to vary with normal stress. Figure 4 shows a plot of temperature increase on the fault (for a given distance of slip) as a function of shear force in the driving piston. Normal force is varied by lowering screws holding up the upper block. It is measured by weighing the force holding up the block for each setting of the lift screws. The coefficient of friction (ratio of shear force to normal force) decreases by increasing normal stress. There is a clear linear relation between shear force and heat generation (for a given slip distance), as would be expected, since no energy is radiated in seismic waves. The coefficient of friction changes with normal stress because at very high stresses the foam rubber lattice fibers that provide the friction (by interlocking) are deformed nonlinearly (bent over) at the interface.

The difference between the results shown in Figure 3 and Figure 4 dramatically illustrates the difference in heat generation and dynamics between the stick-slip and stable sliding models. If the analogy with earthquakes is valid, these results provide an explanation for the heat flow paradox. Further study of the stick-slip process in foam rubber will be required to check that there are no special properties of the foam rubber model that makes the model inappropriate asan analog for earthquake (*e.g.* non-linearities, relaxation effects, adhesion).

## CONCLUSION

The foam rubber model of stick-slip suggests that dynamic normal vibrations or variations in normal stress (with possible fault opening) may play an important role in the stick-slip process. Comparison of the results for frictional heat generation in the stick-slip model with results from another foam rubber model which slides stably without stick slip clearly shows that stick-slip with associated interface separation lowers frictional heat generation, and supports the suggestion that this phenomenon may be an explanation for the heat flow paradox.

## REFERENCES


Anooshehpoor, R., and J.N. Brune, (1992), *Normal Vibration, Fault Separation, and Heat Generation During Stick-slip Events in Foam Rubber*, Seismol. Res. Letts. Vol. 63, no. 1, p. 75.

Barquins, M., and R. Courtel, (1975), *Rubber Friction and Rheology of Viscoelastic Contact*, Wear, vol. 32, 133–150.

Blandford, R. R., (1975), *A Source Theory for Complex Earthquakes*, Bull. Seismol. Soc. Am., vol. 65, 1385–1405.

Brown, S. R., W. A. Olsson, and J. N. Brune, (1991), *Experimental Observation of Interface Separation Waves During Stick-slip*, EOS, vol. 72, no. 44, p. 326.

Brune, J. N., T. L. Henyey, R. F. Roy, (1969), *Heat Flow, Stress, and Rate of Slip along the San Andreas Fault, California*, J. Geophys. Res., Vol. 74, 3821–3827.

Brune, J. N., (1970), *Tectonic Stress and the Spectra of Seismic Shear Waves from Earthquakes*, J. Geophys. Res., vol. 75, 4997–5009.

Brune, J. N., (1976), *The Physics of Earthquake Strong Motion, Seismic Risk and Engineering Decision*, edited by C. Lomnitz and E. Rosenblueth, 140–177.

Brune, J. N., P. A. Johnson, and C. Slater, (1989), *Constitutive Relations for Foam Rubber Stick-slip* Seis. Res. Letts., vol. 60, p. 26.

Brune, J. N., P. A. Johnson, and C. Slater, (1990), *Nucleation, Predictability, and Rupture Mechanism in Foam Rubber Models of Earthquakes*, J. of Himalayan Geol., vol. I(2), 155–166.

Brune, J. N., S. Brown, and P. A. Johnson, (1993), *Rupture Mechanism and Interface Separation in Foam Rubber Models of Earthquakes: A Possible Solution to the Heat Flow Paradox and the Paradox of Large Overthrusts*, Tectonophysics, 218, 59-67.

Byerlee, J. D., (1990), *Friction, Overpressure and Fault Normal Compression*, J. Geophys Res. Letts., 17, 2109–2112.

Byerlee, J.D., (1992), *The Change in Orientation of Subsidiary Shears Near Faults Containing High Pore Fluid Pressure*, Tectonophysics, v.211, 295–303.

Castro, R. R., J. G. Anderson, and J. N. Brune, (1991), *Origin of High P/S Spectral Ratios*

*from the Guerrero Accelerograph Array*, Bull. Seismol. Soc. Am, vol. 81, no. 6, 2268–2288.

Comninou, M., and J. Dundurs, (1977), *Elastic Interface Waves Involving Separation*, Transaction of the ASME J. Appl. Mech., 222–226.

Comninou, M., and J. Dundurs, (1978a), *Can Two Solids Slide Without Slipping?*, J. of Solid Structures, 251–260.

Comninou, M., and J. Dundurs, (1978b), *Elastic Interface Waves and Sliding Between Two Solids*, J. of Applied Mechanics, 325–345.

Dahlen, F. A., (1974), *On the Ratio of P-wave to S-wave Corner Frequencies for Shallow Earthquake Sources*, Bull. Seis. Soc. Am., vol. 64, 1159–1180.

Dieterich, J. H. (1979a), *Modeling of Rock Friction 1. Experimental Results and Constitutive Equations*, J. Geophys. Res., 84, 2161–2168.

Dietericht, J. H., (1979b), *Modeling of Rock Friction 2. Simulation Preseismic Slip*, J. Geophys. Res., 84, 2169–2175.

Dietericht, J. H., (1981), *Constitutive Properties of Faults with Simulated Gouge*, in Mechanical Behavior of Crustal Rocks, edited by N. L. Carter, M. Friedman, J. M. Logan, and D. W. Streans, Geophys. Monogr. Ser., 24, AGU, Washington, D. C., 103–120.

Dieterich, J. H., (1986), *A Model for the Nucleation of Earthquake Slip*, in Maurice Ewing volume, edited by Das, S., J. Boatwright, and C. H. Scholz, Am. Geophys. Union, pp. 37–47.

Freund, L. B., (1978), *Discussion*, ASME J. of Applied Mechanics, vol. 45, 226–228.

Haskell, N. (1964), *Total Energy and Energy Spectral Density of Elastic Wave Radiation from Propagating Fault*, Bull. Seism. Soc., Am. vol. 54, 1811–1841.

Heaton, T. H., (1990), *Evidence for, and Implications of Self-healing Pulse of Slip in Earthquake Rupture*, Phys. Earth. Planet Sci., 16, 1–20.

Hubbert, M. K., and W. W. Rubey, (1959), *Role of Fluid Pressure in Mechanics of Overthrust Faulting-1. Mechanics of Fluid-filled Porous Solids and its Application to Overthrust Faulting.* Geol. Soc. Am. Bull., 70, 115-166.

Lachenbruch, A. H., (1980), *Frictional Heating, Fluid Pressure, and the Resistance to Fault Motion*, J. Geophys. Res., 85, 6079–6112.

Monlar, P., B. E. Tucker, and J. N. Brune, (1973), *Corner Frequencies of P and S Waves and Models of Earthquake Sources*, Bull. Seism. Soc. Am., vol. 63. 2091–2104.

Papageorgiou, A. S., and K. Aki, (1983), *A Specific Barrier Model for the Quantitative Description of Inhomogeneous Faulting and the Prediction of Strong Ground Motion*, Part II. Applications of the Model, Bull. Seism. Soc. Am., vol. 73, 953–978.

Rice, J. R., (1992), *Fault Stress States, Pore Pressure Distributions, and the Weakness of San Andreas Fault*, in Evans, B., and Wong, T. F., eds., Earthquake Mechanics and Transport Properties of Rocks: London, Academic Press, 475–503.

Ruina, A. L., (1980), *Friction Laws and Instabilities: A Quasi-static Analysis of Some Dry Frictional Behavior*, Ph. D. Dissertation, Brown University, Providence, RI.

Schallamach, A., (1971), *How Does Rubber Slide?* Wear, 17, 301–312.

Sleep, N. H., M. L. Blanpied, (1992), *Creep, Compaction, and the Weak Rheology of Major Faults*, Nature, London, 359, p. 687–692.

Smith, K. D., J. N. Brune, and K. F. Priestley, (1991), *The Seismic Spectrum, Radiated Energy and the Savage and Wood Inequality for Complex Earthquakes*, Tectonophysics, 188, 303–320.

Vernon, F., J. N. Brune, and J. Fletcher, (1989), *P/S Spectral Ratios for the Anza Array: Evidence for Normal Motions During Fault Rupture*, Trans. Am. Geophys. Union, 70, p.1226.

Walter, W. R, J. N. Brune, (1993), *Spectra of Seismic Radiation From a Tensile Crack*, J. Geophys. Res., vol. 98, no. B3, 4449-4459.

Zeng, Y., J. N. Brune, J. G. Anderson, (1992), *Dynamic Modeling of Fault Tensile Motion and Opening Wave Propagation*, EOS, vol. 73, No. 43.

# Porosity Loss in the Evolution of Accretionary Wedges: Some Mechanical and Seismic Implications

Dan M. Davis, Department of Earth and Space Sciences
State University of New York at Stony Brook
Stony Brook, NY 11794

From the time of their initial accretion, sediments in accretionary wedges undergo a series of physical and chemical changes that eventually yield sedimentary rocks whose mechanical behavior is vastly different than at the time of their accretion. Although many aspects of these changes are poorly understood, their effects upon the structural geology and the seismicity of forearcs are clearly demonstrable. Three important aspects of these changes have become increasingly evident over the past few years: 1) strain localization as it affects the organization of diffuse deformation into discrete thrusts, 2) strengthening of the sediments with progressive fluid loss as it influences the large-scale structural geology and mechanical evolution of the accretionary mass, and 3) the roles of porosity loss, fluid-related chemical changes, and overpressuring in the determination of the nature (aseismic stable sliding vs. seismic stick-slip) of the deformation.

Newly accreted sediments demonstrate a mixture of degrees of strain localization, with the development of both discrete shear zones and broad non-localized strain. Only a short distance into an accretionary wedge, discrete faulting with fold-and-thrust belt style ramp-flat geometries become common. Because of the importance of hydrologic processes in accretionary wedges and the ability of fractures to dominate fluid flow, it is important to come to a better understanding of the mechanics of both the initial, often discontinuous faulting observed in the trench area and the transition to larger-scale faulting found within much of the wedge itself.

Large-scale deformations in accretionary wedges are governed by the force balance in the sedimentary mass as it overthrusts the plate-boundary basal décollement. Recent mechanical modeling studies indicate that non-uniform variations in strength can lead to the development of distinctive structural styles. In particular, yield strength inhomogeneities due to complex accretionary and hydrologic histories can be important in controlling the vergence of thrust faults, the subduction or underplating of sediments, and the overall geometry of the forearc wedge. It is not clear how these strength contrasts are related to the transitions in slip mode that control the degree of localization of faulting.

The frontal regions of sediment-rich forearcs are invariably aseismic, but at some of the largest known earthquakes have occurred just arcwardand down-dip of that region. At some point within the accretionary wedge there must occur a transition in the mechanical behavior of the sediments from conditions under which any frictional slip that occurs is stable to seismic, stick-slip behavior capable of producing great earthquakes. There are reasons to believe that this transition is related to the processes of fluid and porosity loss. However, it is not clear what, if any, is its relationship to the non-localized/localized strain transition. Nor is it clear how this transition from aseismic to seismic behavior at the 'seismic front' is related to the structurally important strength contrasts that appear to be important in controlling the geometries of outer-arc highs and forearc basins.

These transitions in the localization, yielding, and seismic behavior of accreted sediments are all part of the evolution of a sedimentary packet from its initial accretion to its burial to the deeper parts of the accretionary forearc. Therefore, it is important to recognize the central role played by the progressive loss of fluids in determining the locations in space and time of each of these behaviors. However, because each of these behavioral transitions may occur at a distinct point in that evolition and manifests itself in very different ways, each must also be investigated

individually. It makes sense to begin by reviewing the role of compaction-related yield strength gradients in determining the large-scale structural style of a forearc.

Finite-element and laboratory scale modeling done (Malavielle, 1984; Byrne et al., 1988, 1933) indicate that a distinct gradient in yield strength at depth in a forearc can lead to a reversal in thrust vergence and the development of a structurally and topographically elevated 'pop-up' zone. Above a strong 'backstop', this is a common mechanism for the development of an outer-arc high and a forearc basin behind it that remains largely undeformed. It may also produce less dramatic structural features if the yield strength contrast is less distinct.

Bangs et al. [1990] have found distinct velocity gradients in the forearc of the Lesser Antilles. Regions bounded by two such major zones of high velocity contrast are indicated in the schematic cross-section of Figure 1. The darkly stippled region at left in this figure is the high-velocity backstop, believed to be the relatively strong basement of the volcanic arc. This backstop is responsible for the growth of the outer-arc high on which is located the island of Barbados. It also permits the development of a largely undeformed forearc basin which is 'insulated' by the backstp from the prevalent contractional deformation of the forearc. The lightly stippled region in the center is a region within the accretionary wedge that appears to have moderately high seismic velocities compared to the surrounding sediments. The broad structural arch over the toe of this area and the distinct backthrust there are as one would predict for a weak backstop.

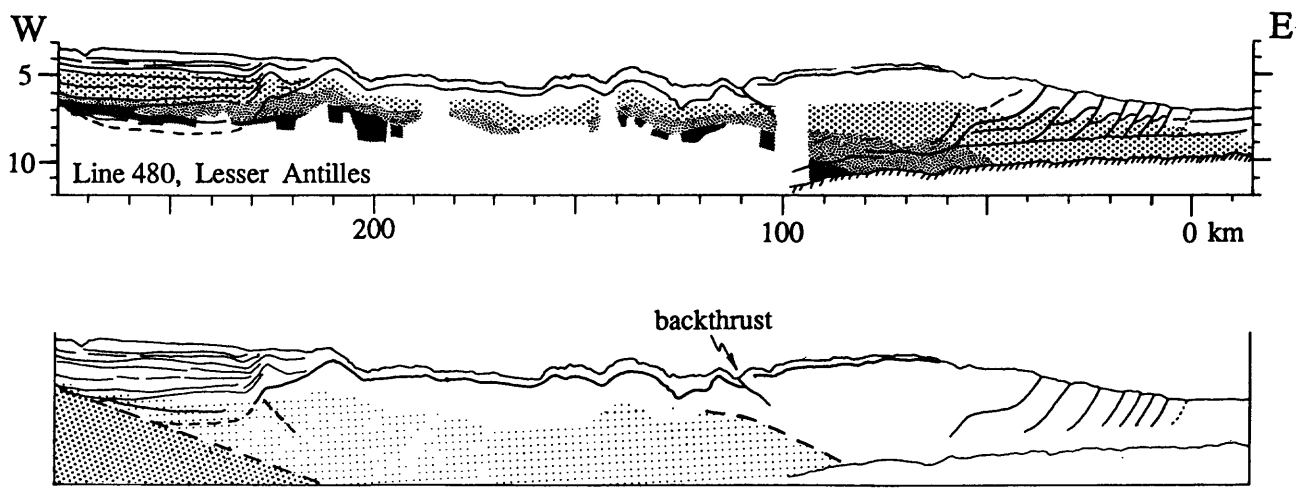



*Figure 1) Seismic velocities in the Southern Lesser Antilles accretionary wedge [Bangs et al., 1990]. In the figure at top, the four shadings (white to dark) represent 0.5 km/sec velocity intervals: white represents velocities less than 2.5 km/sec and the darkest shade indicates velocities over 3.5 km/sec. In the lower figure, the seismically fastest region of the accretionary wedge (and presumably the strongest as well) is indicated by light stippling. Above and at the front of this area there is a structurally elevated zone and a local vergence reversal, as predicted by mechanical models. The high-velocity main backstop, thought to consist of volcanic arc rocks, is indicated schematically by the heavily stippled triangle at left, beneath the forearc basin.*

The process of accretion is thought to be discontinuous, with large fluctuations in the rate and lithology of sediment input. This, combined with the time-dependence of the processes of dewatering and lithification, means that a hiatus or marked lithologic change in accretion at some point in the history of the forearc might easily produce a lasting region of significantly lower porosity, higher degree of lithification, and higher yield strength. To evaluate this possibility and its structural implications it is necessary to know more about how accreted sediments become progressively stronger.

It has been appreciated for some time that in addition to the chemical processes surrounding clay dehydration and the formation of cements, porosity loss is one of the factors controlling the evolution through time of the yield strengths of sediments. Laboratory studies have suggested that strength contrasts large enough to produce significant structural effects can result from porosity contrasts of only a few percent. Observations such as those described above indicate the significant effect that inhomogeneities in the dewatering and lithification processes can have on the large-scale structural geology of a forearc wedge. They lead us inevitably to ask what might be the effects of these processes on the behavior of the shear zones themselves, particularly in terms of the degree to which they show discrete slip on faults and whether those faults slip seismically or aseismically.

The main plate-boundary thrust at subduction zones produced many of the largest and most damaging of earthquakes. Byrne et al. (1988) have shown that a wide variety of forearcs share similar patterns in their distributions of seismicity, with several tens of kilometers of the plate and plate boundary nearest the trench apparently aseismic. The maximum moment for an earthquake at most margins is approximately proportional to the cube of the seismogenic zone width, so an appreciation of the factors controlling that width is important in understanding seismic risk.

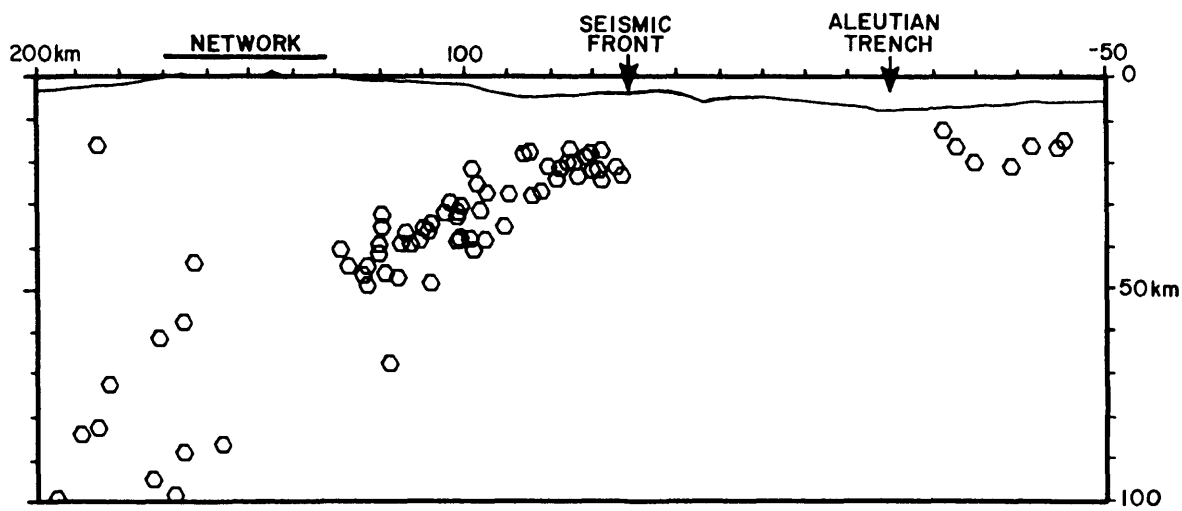



*Figure 2) A cross section showing hypocenters of earthquakes detected by the Adak seismic network and relocated by Engdahl and Gubbins [1987]. Note the aseismicity of the accretionary wedge area, between the flexural events seaward of the trench and the onset on upper-plate and plate-boundary seismicity at a boundary that we have called the 'seismic front'.*

The cross-section in Figure 2 illustrates the typically aseismic character of accretionary wedges. Although the local seismic network was able to detect flexural events past the trench and numerous events arcward of the 'seismic front', the accretionary wedge in this (as in numerous other examples around the world) appears to be aseismic.

The prevalence of aseismicity in sediment-rich accretionary wedges around the world suggests that the aseismic behavior of the frontal regions of forearcs is related to the physical state of the sediments common to that setting. The association of aseismic behavior with porous and

relatively weak sediments is reinforced by the observation that backstops capable of producing pronounced outer-arc highs and forearc basins are usually seismic while the seismically slower and more porous sediments trenchward of them are invariably aseismic.

The role of the high-temperature transition from seismic to aseismic slip at the deep end of the seismogenic zone at plate boundaries is well appreciated, but the nature of the shallow, trenchward limit to seismicity is less clear. In order to really understand what controls the level of seismic risk at a subduction zone we must learn more about the transformation of highly porous accreted sediments into material that can localize strain on discrete faults and eventually store up elastic strain energy, releasing it in frictional stick-slip events along those faults. To this end, it is necessary to consider what can be learned from laboratory studies of the modes of failure in porous sedimentary rocks.

Recent work (Zhang et al., 1993a,b) has allowed the identification of a boundary between non-localized, non-dilatant, 'ductile' behavior and shear-localized, dilatant 'brittle' behavior in porous sediments as a function of two parameters known to vary rapidly across accretionary wedges: porosity and effective pressure. Having precompacted sediments to a variety of initial porosities, we subjected them to triaxial deformation in order to determine whether they behave in a macroscopically 'brittle' or 'ductile' manner. Combined with data from other studies, this leads to a distinct boundary between these behaviors in each sedimentary rocks studied. When sediments from the ODP Nankai drilling leg were studied (Zhang et al., 1993b), it was found that the 'brittle'-'ductile' transition, as described above, occurs at a porosity ranging from about 40% to perhaps 20%, depending upon the effective confining pressure. At high porosity (conditions likely to be found near the toe of an accretionary wedge) the behavior is 'ductile'.

The large dots in Figure 3 represent experimental determinations of behavior and the small dots correspond to the in-situ conditions of sediments recovered during drilling. To do this, we have made the conservative assumptions that porosity is conserved in drilling (unlikely at very

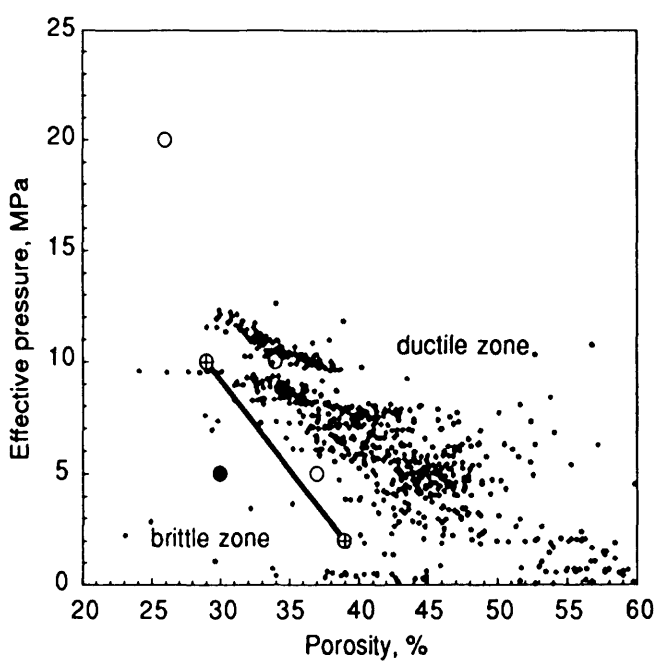



*Figure 3) A brittle-ductile transition map based upon triaxial experiments on ODP Nankai leg samples (large circles). An open circle indicated "ductile' behavior, a circle wit a cross showedtransitional behavior, and a solid circle indicates a sample that deformed in a macroscopically 'brittle' manner. The small sots represent in-situ porosity and effective pressures, making the assumptions described in the text. The shallowest samples (those with the lowest effective pressures) are most disrupted by drilling and should probably be ignored.*

*Note that most of the toe of the Nankai wedge, with the marked exception of a major thrust zone apparently falls into the 'ductile', non-localizing regime. Elevated fluid pressures may push more points downward in effective pressure and into the 'brittle' region.*

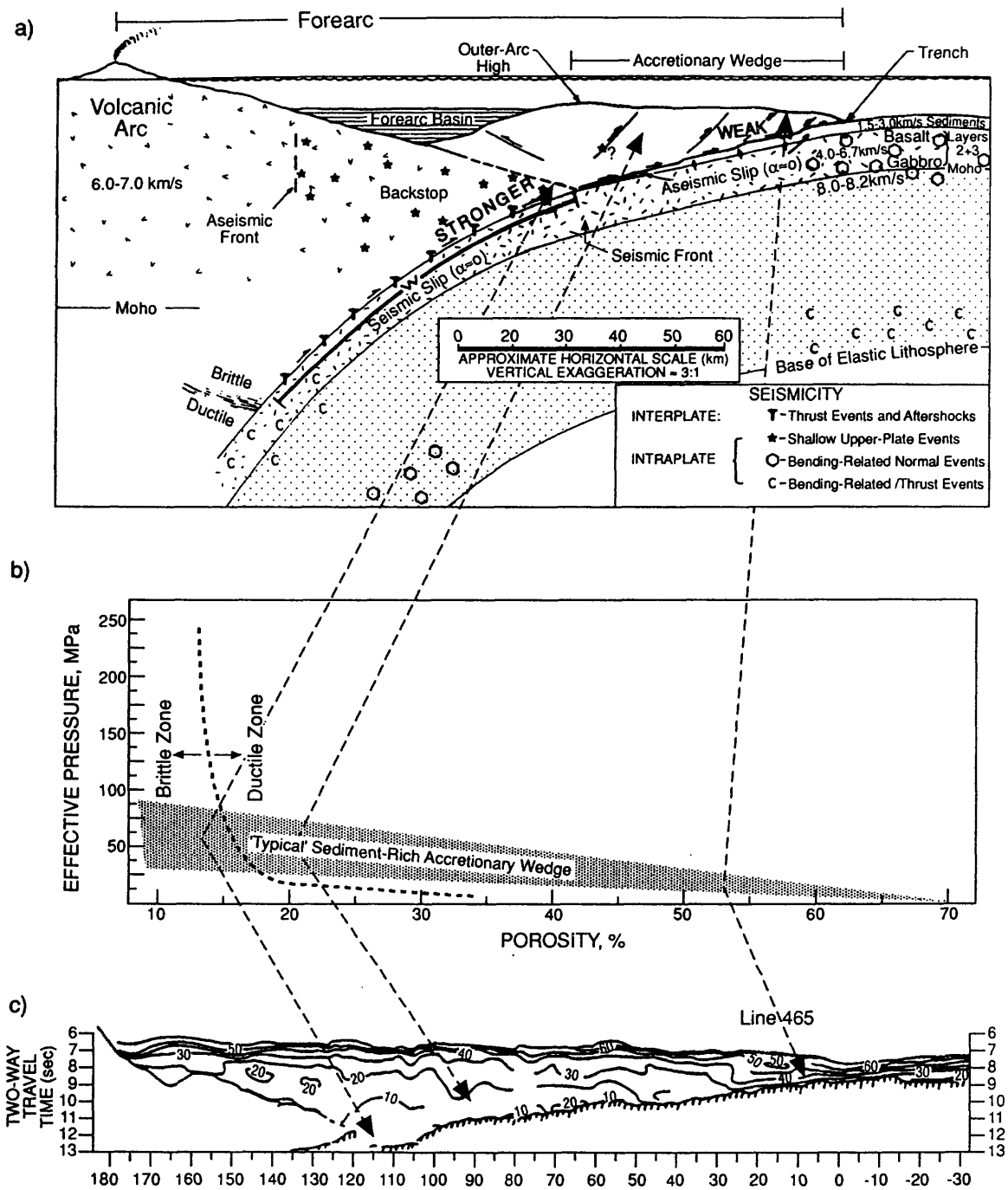



*Figure 4) A schematic comparison among:*

*a) a cross-section through a forearc, illustrating the relationship between a low-porosity, strong backstop and the outer-arc high, forearc basin, and the structural vergence,*

*b) a generalized version of a failure mode transition map such as that in Figure 4, and*

*c) the inferred porosity profile for the Lesser Antilles forearc accretionary wedge, as determined by Bangs et al. [1990]. The aseismicity and structural weakness of most of the accretionary wedge may simply result from the high porosity and low degree of compaction of the sediments there.*

*Figure after Zhang et al. (1993a).*

shallow depths) and that fluid pressures are hydrostatic (which may lead to substantial local overestimates of the effective pressure). This figure suggests that localization of shear in the frontal region of accretionary wedges my be a complicated process involving fluid loss and overpressuring. If the initial conditions do not permit the formation of highly discrete shear, then protothrusts in trenches may result from localized, transitory increases in fluid pressure sufficient to reduce the effective pressure enough to allow a transitional, partial localization of shear. That in turn may affect the permeability by enough to permit the formation of weakly localized shear zones. It is interesting to note that steeply-dipping, discontinuous shear zones are observed in seismic reflection lines in the protothrust zone just seaward of the frontal thrust in at least some accretionary wedges, including those at the Oregon and Nankai margins.

In summary, the loss of fluid from pore spaces and the development of high fluid pressures appear to be the common factors linking several distinct aspects of the tectonics of forearcs. These factors appear to control the initial ability of strain to become concentrated on discrete faults. The transition to fully shear-localized, dilatant behavior is a necessary (but not sufficient) condition for the onset of seismic slip. Some other factors that may contribute to that causing that transition, such as dewatering reactions and cementation are also intimately tied to the flow of fluids within the forearc. Finally, the variations of yield strength created by these same fluid-controlled processes of compaction and cementation exert an important control on the overall shape of and strain distribution within the forearc, as well as on the magnitudes and vergences of slip on thrusts.

## REFERENCES


Bangs, N. L. B., G. K. Westbrook, J. W. Ladd, and P. Buhl: Seismic velocities from the Barbados Ridge: Indicators of high pore fluid pressures in an accretionary complex, *J. Geophys. Res., 95*, 8767-8782, 1990.

Byrne, D., D. M. Davis, and L. Sykes: Loci and maximum size of thrust earthquakes and the mechanics of the shallow region of subduction zones. *Tectonics, 7*, 833-857, 1988.

Byrne, D. E., W.-H. Wang, and D. M. Davis: Mechanical role of backstops in the growth of forearcs, *Tectonics, 12*, 123-144, 1993.

Engdahl, E.R., and R. Gubbing, Simultaneous travel-time inversion for earthquake location and subduction zone structure in the Central Aleutian Islands, *J. Geophys. Res., 92*, 13,855-13,862, 1987.

Malavieille, J.: Modelization experimentale des chevauchements imbriqués: Application aux chaines de mountagnes, *Bull. Soc. Geol. France, 7*, 129-138, 1984.

Silver, E. A., M. J. Ellis, N. A. Breen, and T. H. Shipley: Comments on the growth of accretionary wedges, *Geology, 13*, 6-9, 1985.

Zhang, J., D.M. Davis, and T.-f. Wong: Brittle-ductile transition in porous sedimentary rocks: geological implications for accretionary wedge aseismicity, *J. Structural Geol., 15*, 7, 819-830, 1993a.

Zhang, J., D.M. Davis, and T.-f. Wong: Failure modes of tuff samples from Leg 131 in the Nankai accretionary wedge, in Hill, I.A., Taira, A., Firth, J.V. et al., *Proceedings of the Ocean Drilling Program, Scientific Results, Vol. 131*, 275-281, 1993b.

# The Impact of Earthquakes on Fluids in the Crust

R. MUIR WOOD[1] & G.C.P. KING[2]

1 EQE, Newbridge House, Clapton. Glos. GL54 2LG, UK
2 Institut de Physique du Globe, 5 rue René Descartes, 67084 Strasbourg, France.

## Abstract

The character of the hydrological changes that follow major earthquakes has been investigated and found to be critically dependent on the style of fault displacement. In areas where fracture-flow in the crystalline crust communicates uninterrupted to the surface the most significant response is found to accompany major normal fault earthquakes. Increases in spring and river discharges peak a few days after the earthquake and typically excess flow is sustained for a period of 4-12 months. Rainfall equivalent discharges, have been found to exceed 100mm close to the fault and remain above 10mm at distances greater than 50km. The total volume of water released in two M7 normal fault earthquakes in western USA was 0.3-0.5 $km^3$. In contrast, hydrological changes accompanying reverse fault earthquakes are either undetected or else involve falls in well-levels and spring-flows.

The magnitude and distribution of the water-discharge for these events is compared with deformation models calibrated from seismic and geodetic information, and found to correlate with the crustal volume strain down to a depth of at least 5km. Such relatively rapid drainage is only possible if the fluid was formerly contained in high aspect ratio fissures interconnected throughout much of the seismogenic upper crust. The rise and decay times of the discharge are shown to be critically dependent on crack widths, for which the 'characteristic' or dominant cracks cannot be wider than 0.03mm.

These results suggest that fluid-filled cracks are ubiquitous throughout the brittle continental crust, and that these cracks open and close through the earthquake cycle. Seismohydraulic fluid flows have major implications for our understanding of the mechanical and chemical behaviour of crustal rocks, of the tectonic controls of fluid flow associated with petroleum migration, hydrothermal mineralisation and a significant hazard for underground waste disposal.

## Introduction

A clear relationship can be demonstrated between the style of fault displacement and its hydrological 'signature' (Muir Wood and Woo, 1992; Muir Wood and King, 1993). Coseismic strain models can be shown broadly to explain the difference between the hydrological signatures of normal and reverse fault earthquakes, the geographical extent of the hydrological response and even, in general terms, the magnitude of the water-release.

In a region undergoing extensional faulting, continuing strain distributed through the crust causes appropriately oriented fractures to dilate, thereby increasing crustal porosity. Pore-pressures are sustained by slow infilling of the dilated fractures with water from the surface (see Fig. 1). At the time of a major normal fault rupture, strain formerly distributed through the crust becomes concentrated on the fault, and the surrounding crust undergoes elastic rebound in compression. In contrast, in a region undergoing compressional tectonic deformation, negative strain (volume decrease) in

the crust closes fractures and reduces crustal porosity in the interseismic period. At the time of the fault rupture, as strain is transferred into fault displacement the surrounding crust undergoes elastic rebound in extension.

Hence in the region surrounding a normal fault rupture it is expected that the decrease in crustal porosity leads to the expulsion of water following an earthquake. In the region surrounding a compressional fault rupture reduced hydraulic pressures should lead to water being drawn into the crustal volume. In both normal and reverse fault earthquakes the impact on the height of the water-table (and consequent well-levels and superficially sourced spring flows) will be dependent on how effectively this water-table is connected to the fracture-flow system at depth. Where the crystalline rocks of the crust communicate with the surface uninterrupted by the presence of any overlying impermeable sedimentary cover the superficial groundwater should respond to underlying changes in fluid pressures.

River flows provide the most important resource for quantifying these post-seismic changes. By sampling changes in discharge across catchments typically >100km$^2$ in area, it becomes possible to average the crust on a scale whose lateral dimension is of the same order as the hydraulically-conductive thickness of the crust. River discharge data can then be normalised for the area of the drainage basin to achieve a 'rainfall equivalent' discharge either in the form of velocity (for a daily average) or cumulative linear 'thickness'.

**Normal vs. Reverse faults**

We have documented changes in near-surface hydrology following a number of earthquakes in diverse tectonic environments. The form of these hydrological changes can be illustrated from two major high-angle dip-slip earthquakes: the 1959 M7.3 normal fault Hebgen Lake earthquake and the M7.2 1896 reverse fault Rikuu (North Honshu) earthquake. (Earthquakes associated with a significant strike-slip component of displacement reveal more complex hydrological signatures: Muir Wood and King, 1993).

Three river flow profiles (with entirely separate catchments) are shown in Figure 2 from the region of south-east Montana for the period before and after the August 18th, 1959 Hebgen Lake earthquake (data from USGS Water Supply Reports). For both sets of plots average monthly flows are shown for comparison. From the left-hand plots it can be seen that in all three rivers a peak in flow arrived within four days following the earthquake. There was no precipitation around this time to explain such a surge and the increase in flow that follows the earthquake can be seen to have been sustained relative to the trend of the monthly averaged flow curve for more than 60 days. From expected flow curves for these rivers (calibrated from river flows within the same region but outside the influence of the earthquake) it is possible to assess the overall volume of the water release in each catchment that apears to show an almost linear decay following the initial peak (Muir Wood and King, 1993). Typical decay times to half peak flow are 100-150 days. These cumulative flows, normalised for the individual drainage basins around the fault in terms of rainfall equivalent discharges, are shown in Figure 3. The individual catchments with more than a total of 1mm excess discharge have been shaded in three tones corresponding to1-20mm, 21-40mm and >40mm. the total volume of water discharged across all the surrounding catchments is equivalent to ca. 0.5km$^3$.

The Rikuu earthquake of August 31st 1896 involved surface reverse fault rupture, over a distance of 36km with a maximum uplift of 3.5m. This was accompanied by

an antithetic reverse fault rupture located some 15km to the east and in the hangingwall of the main fault. Hence all the near-surface faulting appears to have been reverse implying that all near-surface elastic rebound was extensional. No changes in river flows were reported following the earthquake, but hot springs supplying bath-houses at Oshuku, Tsunagi and Osawa dried up after the earthquake, while there was a significant reduction in flow at the thermal springs at Namari and Yuda (see Fig. 4). All of these lie in the hangingwall of the main fault, to the east and within 20km of the principal antithetic fault. In contrast, close to the northern end of the main fault between the main fault and its antithetic a new hot spring formed at Sengan-Toge following the earthquake.

**Coseismic strain models**

In order to examine in more detail the expected magnitude and extent of strain induced hydrological effects, strain models of coseismic deformation have been generated using a boundary element programme in which a dislocation element is introduced into an elastic medium. Figure 5 is an illustration of the dilational strain changes that accompany a normal fault displacement, both in cross-section (Fig. 5a) and a plan-view approximation (Fig. 5b).The predominant strain changes are compressional. Strain changes of reverse fault earthquakes are largely equivalent although of opposite sign to those of normal fault earthquakes.

**Percolation depth**

In the case of the Hebgen Lake earthquake the observed discharge for a traverse perpendicular to the strike of each of the faults has been compared with predicted discharge for a two dimensional coseismic strain model. For the Hebgen Lake earthquake the closest fit is with the 5km depth prediction both in general shape and in amplitude. The fit with the data is remarkably good and the data is not very sensitive to where the profile is taken.

This comparison suggests that all the water that appears at the surface appears to be associated with fracture systems extending to considerable depth. A large part of the coseismic volumetric strain can be shown to be associated with the closure of cracks and expulsion of water.

**Characteristic fracture apertures**

The most straightforward way to describe the form of the observed flows is to assume that they result from a series of planar uniform cracks open at the surface and closed at some depth and subject to a pressure change over a depth range (Muir Wood and King, 1993). The crack width is found to be the parameter most strongly controlling the decay of the flow vs. time profile while the dead depth (depth to the top of the section of the crack undergoing a sudden strain) and the crack width together determine the rise time. Most of the observations are fit by models with crack widths of about 0.03mm and effective dead depths of about 2km.

To obtain cumulative rainfall equivalent discharges of 20-40mm implies characteristic crack separations of less than 4m. The similarity that exists among the duration and rise-time of hydrological signatures in a wide variety of earthquakes suggests that this property is widespread.

**Underground waste disposal.**

Hydrological changes that accompany earthquakes may significantly alter the ground

water regime, potentially causing fractures to reverse their flow, or stagnant bodies of water to become involved in a sudden movement towards the surface. Where underground repositories for chemical or nuclear waste are constructed with the intention of relying on advantageous ground water flow, either in terms of rates or directions of fluid transport, coseismic changes in flow may have a deleterious impact, in the most extreme cases causing contaminated water to pass rapidly to the surface. Fluid flow measured in the interseismic period is likely to be entirely distinct to that which prevails immediately following the earthquake.Seismohydraulic processes need to be given particular consideration when planning the underground disposal of hazardous wastes.

**Mineralisation**

The fundamental differences in strain cycling between extensional and compressional tectonic environments are likely to have a profound impact on the conditions and styles of mineralisation. In the extensional environment interseismic crustal dilation leads to an increase in crustal porosities which may be topped up from the surface water table. Repeated flushing of water through the top five kilometres of the crust, over the 1000 or more seismic cycles that make up the tectonic activity of a rift bounding fault, will significantly affect rock chemistry. For trace ionic species disseminated through the rockmass, reprecipitation to form concentrated vein mineral deposits is most likely to occur in the immediate aftermath of an earthquake when the watermass is squeezed to shallower and cooler levels in the crust.

In contrast, in a compressional tectonic environment fluid is squeezed out between earthquakes, producing such phenomena as mud-volcanoes. The rapid recharge of the dilated rockmass following a reverse fault displacement may be accomplished in part by the inflow of water from deep-sourced hydrothermal fluids already saturated with dissolved ionic species. Mineralisation is therefore most likely in the interseismic period.

**Seismohydraulics and petroleum migration**

Over a significant proportion of the crust there are extensive sedimentary cover sequences of very low permeability, that prevent, at least in the short-term, any connection between basement fluid pressures and near-surface aquifers. If increased fluid-pressures cannot be relieved through flow to the surface, then fluid flow will be lateral, either within the fractured crust, or more often within some overlying aquifer towards the base of the sedimentary cover.

Following a normal fault displacement fluids will tend to migrate away from the vicinity of the fault. The volumes of displaced fluid can be predicted to be comparable to those found emerging at the surface where there is no impermeable cover (ie. ca. $0.5km^3$ of fluid migrating laterally from a region $10,000km^2$, in area). The flux on the margins of such an area, entirely covered by an impermeable overburden would be more than $1,000m^3/m$ of perimeter: in a 10m thick aquifer with a 10% dynamic porosity this would move the fluid 1km. Following a reverse fault displacement fluid flow would be drawn towards the fault. Such pulsed episodes of seismohydraulic fluid flow are likely to have a significant impact on the migration of hydrocarbons, in particular when tectonic activity is resumed fairly late in the evolution of a sedimentary basin.

The flows created by the large-scale and sudden changes in pore-pressures accompanying seismic strain-cycling can oppose the prevailing fluid flow regimes created by gravitational or thermally-driven fluid flows. Tectonics could therefore

drive fluids into reservoirs that might otherwise be discounted as potential drilling targets. Employing coseismic and interseismic strain models it is possible to predict the magnitude and extent of these multiple episodes of fluid flow as part of basin analysis. Coseismic and interseismic strain changes may also explain the development of zones of over- and under-pressurisation within confined fluid reservoirs.

## References

MUIR WOOD R. & KING G.C.P. 1993. Hydrological signatures of earthquake strain. Journal of Geophysical Research, in press.

MUIR WOOD R. & WOO G. 1992. Tectonic Hazards for Nuclear Waste Repositories in the UK, Report by YARD Ltd for HMIP, UK Department of Environment, No. PECD 7/9/465.

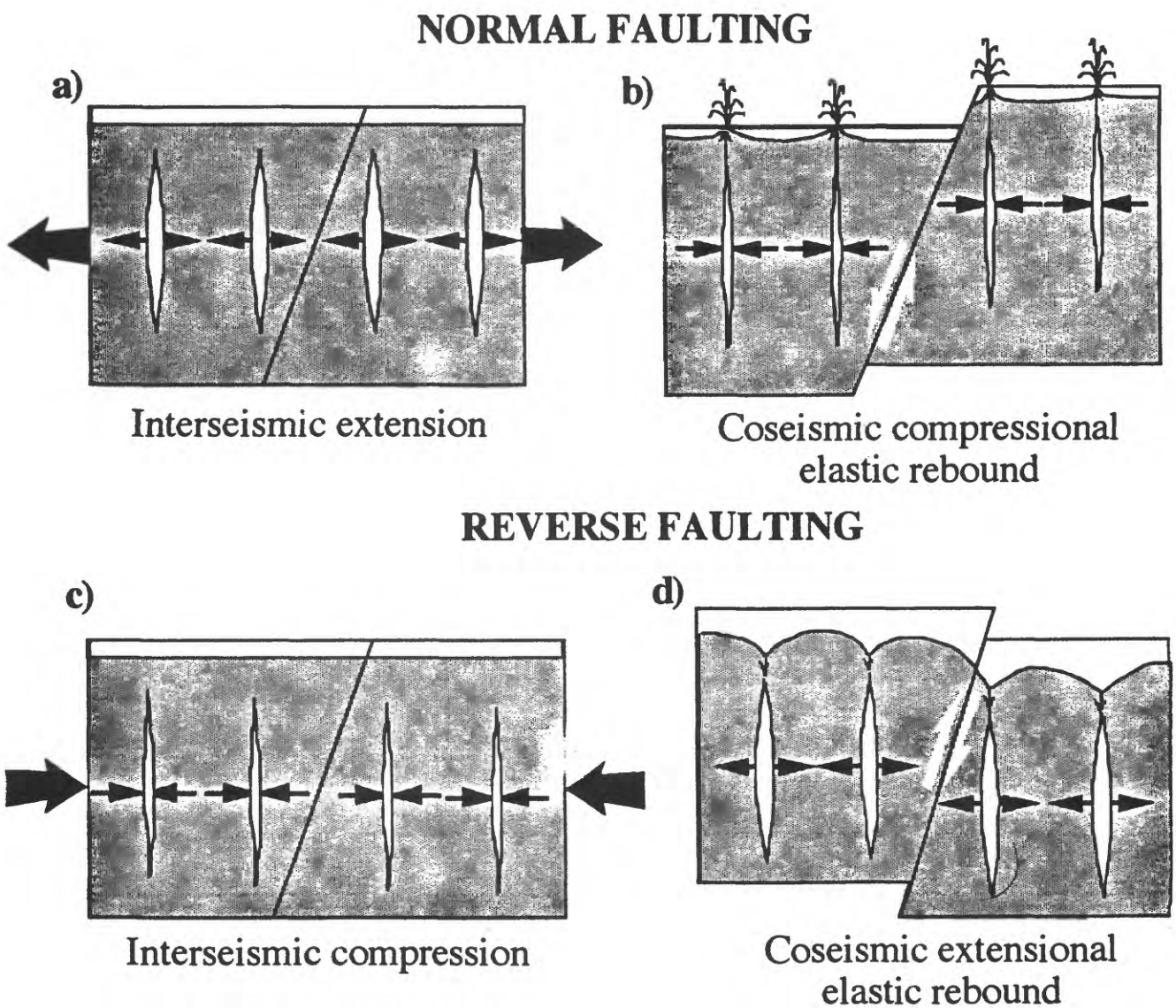



Figure 1. For extensional faulting the interseismic period a) is associated with crack opening and increase of effective porosity. At the time of the earthquake b) cracks close and water is expelled. For compressional faulting the interseismic period c) is associated with crack closure and the expulsion of water. At the time of the earthquake d) cracks open and water is drawn in.

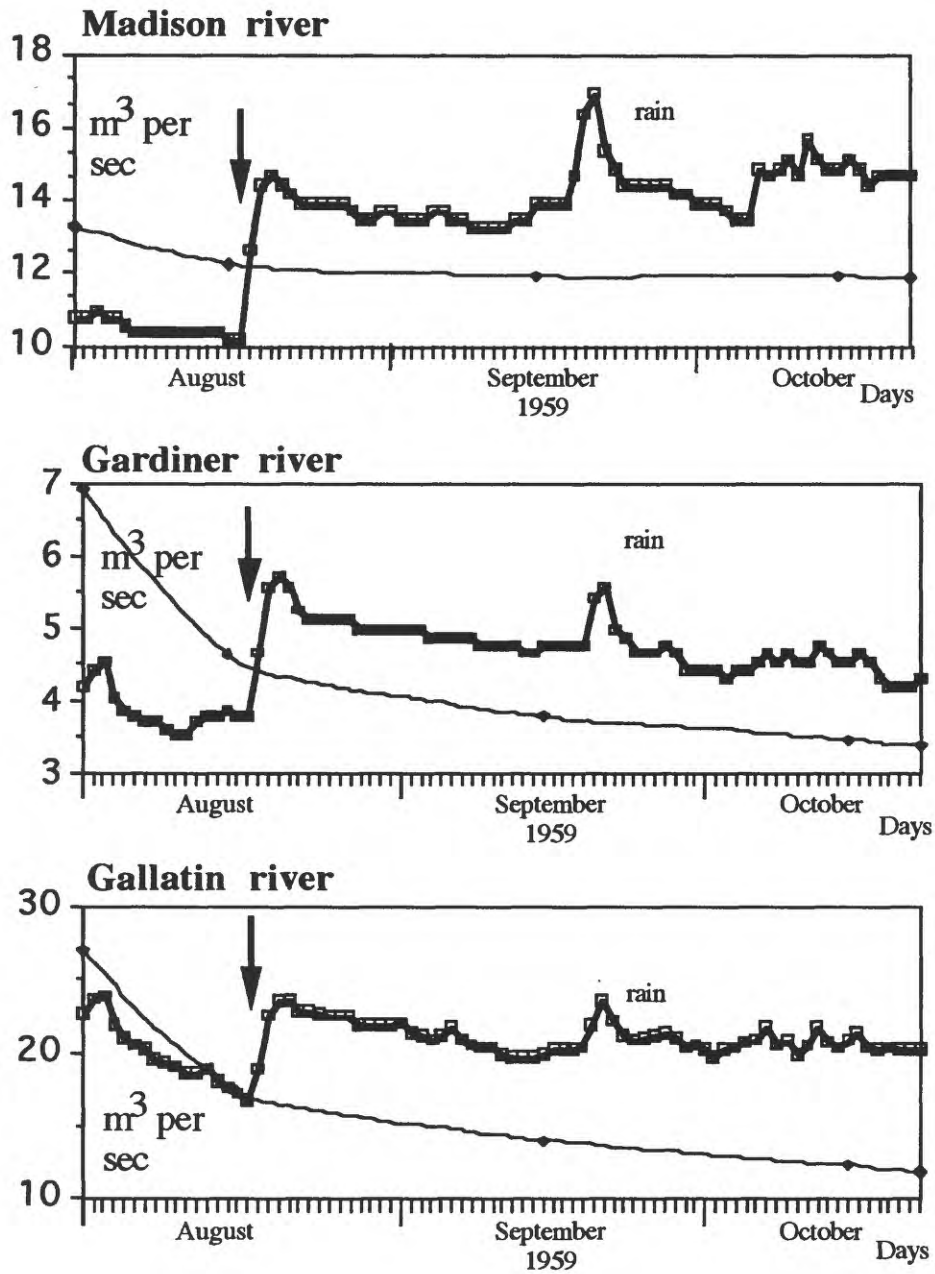



Figure 2. River flow data for three rivers in the vicinity of the 1959, Aug. 17th Hebgen Lake, Montana earthquake, in the days around the time of the earthquake. Average monthly flows plotted as thin line.

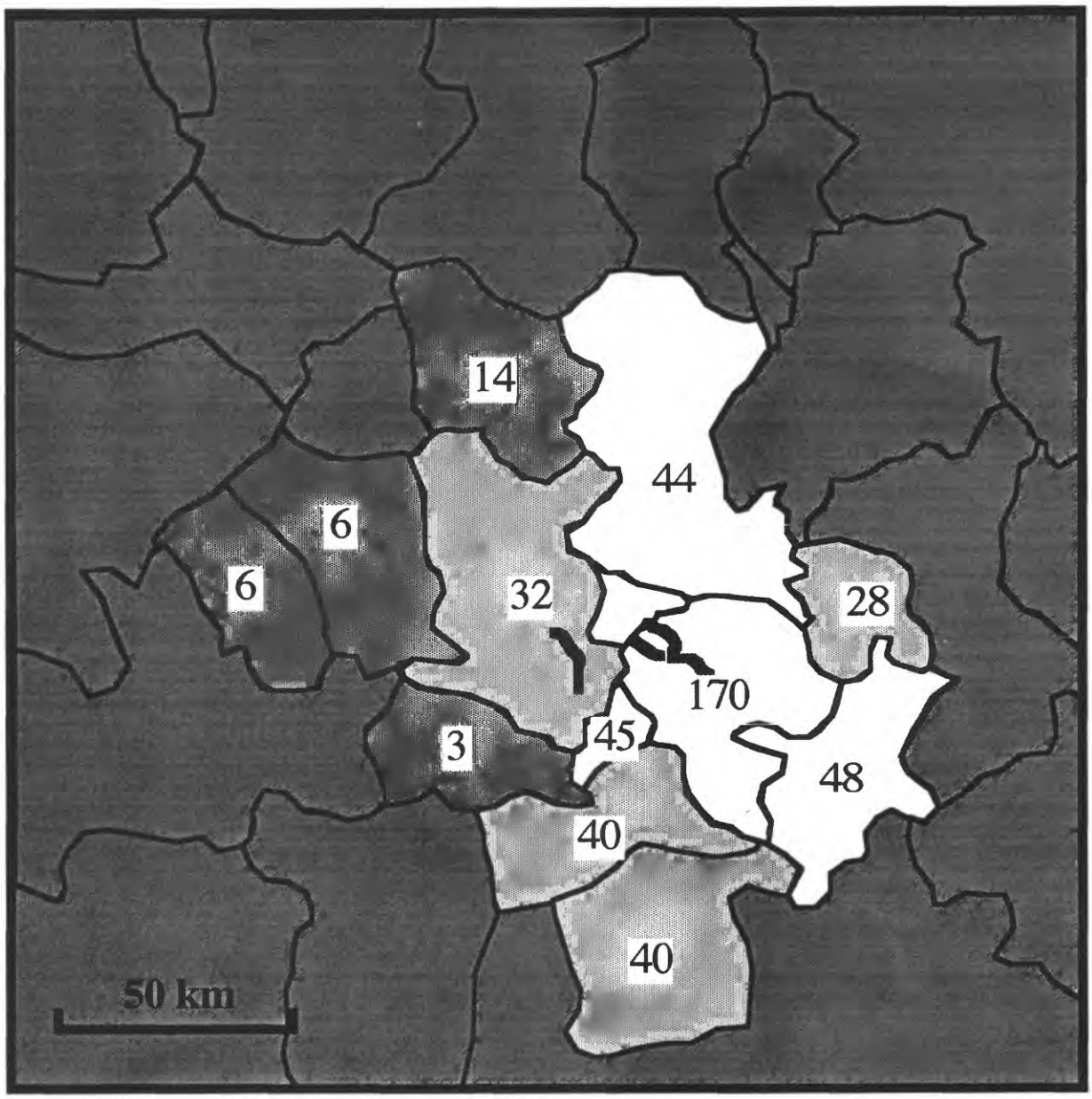



Figure 3.Cumulative excess flow following the Hebgen Lake earthquake normalised for the area of the catchments into excess rainfall equivalent (mm). Mid-grey shading indicates catchments with less than a 1mm excess flow; three lighter tones correspond to 1-20mm, 20-40mm and >40mm.

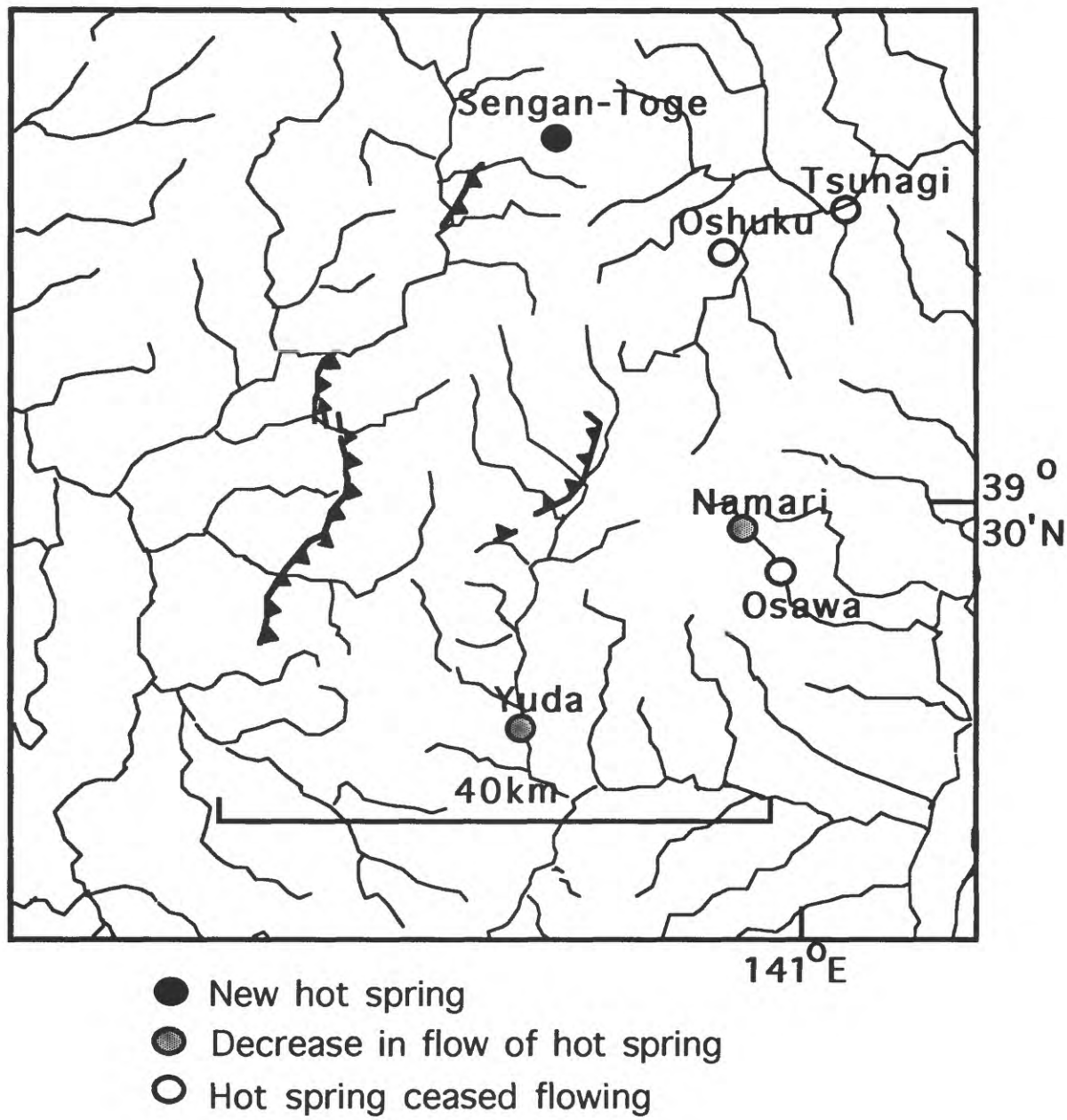



Figure 4. Impact on hot-springs of the August 31st 1896 Rikuu, (North Honshu) M7.2 reverse fault earthquake. Barbed lines are surface-fault ruptures; barbs on upthrown side of the fault.

# Dip-slip motion on a Dipping Fault

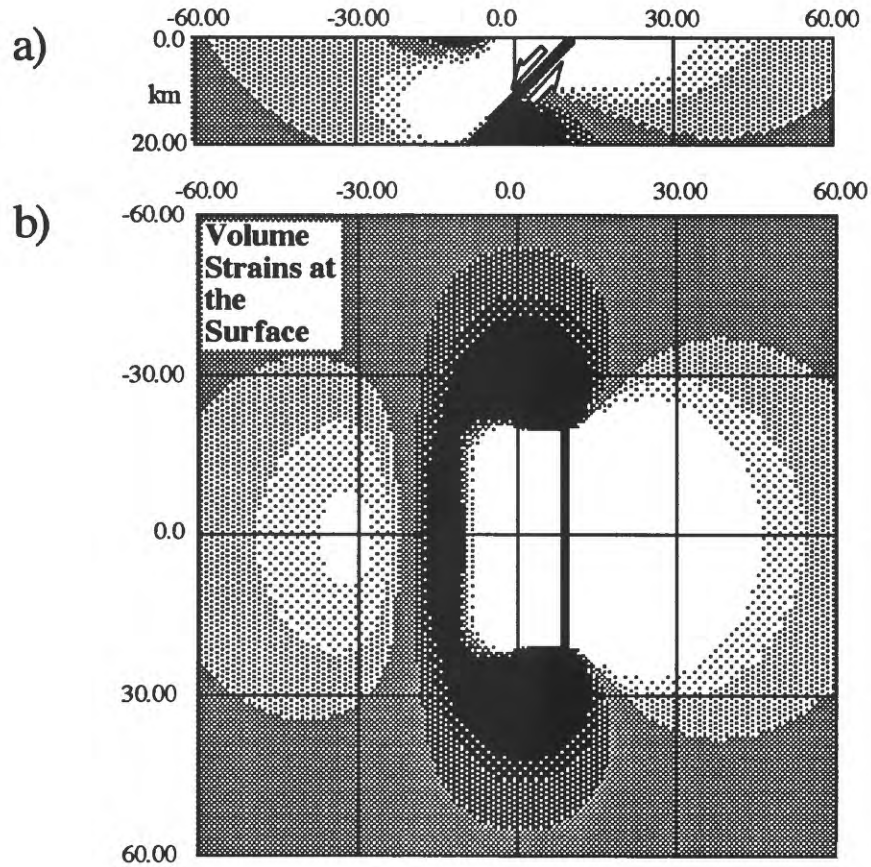



Figure 5. Cross-section (a) and plan-view (b) of strains modelled around a normal fault. The scale units are kilometers, slip for (a) is 1 metre. Strain levels are shaded from a background grey, to lighter shades (negative strain) and to darker shades (positive strain): strain-steps of 2 x 10-5.

# Long-Lived Permeability Changes In The Shallow Crust Associated With Earthquakes

S. Rojstaczer
Duke University, Dept. of Geology, Box 90230, Durham, NC 27708

S. Wolf
U.S. Geological Survey, M.S. 988, Menlo Park, CA 94025

R. Michel
U.S. Geological Survey, M.S. 431, Reston, VA 22092

*Abstract*

We performed reconnaissance of the groundwater and surface water of an earthquake impacted drainage basin to place bounds on the time constant associated with earthquake-caused permeability increases. Groundwater flow rates are high (on the order of 100s of meters per year) and residence times are short. Three years after the earthquake, stream chemistry is still elevated over pre-earthquake levels suggesting that permeability increases associated with earthquakes can be long lived.

## Introduction

Understanding the nature of seismically induced permeability changes yields some fundamental insight into the evolution of groundwater flow in areas of crustal unrest. Such understanding also has practical implications: studies of shallow groundwater-faulting interaction provide a way of assessing the possibility of seismically induced flooding at the proposed Yucca Mountain high-level nuclear waste respository (1). In response to the Loma Prieta earthquake (10/17/89, M 7.0), pervasive changes occurred in the hydrology of the San Lorenzo drainage basin (Fig. 1). Stream flow, spring flow, stream chemistry and water table levels were all impacted by the Loma Prieta earthquake (2). The hydrologic response of the San Lorenzo drainage basin - rapid increases in stream flow and decreases in water levels in wells - can be explained if permeability was greatly enhanced by the Loma Prieta earthquake throughout the drainage basin. Other commonly cited mechanisms to explain earthquake induced hydrologic changes cannot adequately explain the observations (3). In this paper we examine the time constant associated with these permeability changes. We describe hydrologic reconnaissance performed in 1992 and 1993 in order to assess the state of groundwater flow in the drainage basin and the changes that have taken place in the hydrology of the area relative to historical data.

## Residence Time Of Groundwater

We sampled and hydraulically tested wells in the region earlier documented to have undergone significant drops in water table elevation (Fig. 2). These wells tap Tertiary marine sandstones and shales that are present throughout much of the basin and the region. Chemical analysis of these wells indicates that while these are generally bicarbonate waters, sulfate waters are present in some wells that tap the shales. Water in the wells is oversaturated with respect to calcite, with the saturation index typically about 1.0.

Hydraulic testing of these wells suggests that the hydraulic conductivity of the formations is quite high. The hydraulic conductivity is generally between $2 \times 10^{-4}$ and $5 \times 10^{-4}$ cm/s. The specific yield (drainable porosity) is typically about 0.10-015. These

values indicate that the matrix has significant permeability. The high values for hydraulic conductivity suggest that residence time for groundwater in the basin is short. For example, given an average hydraulic gradient of 0.05 and typical values of hydraulic conductivity and specific yield (assuming that drainable porosity is not significantly different than total porosity), the time it takes for groundwater to enter the system as recharge and exit as streamflow is on the order of 1-5 years (4).

Tritium analyses of well water and spring water sampled in 1992 also suggest that groundwater residence time is generally very short (Fig. 3). High levels of tritium (half life of 12.43 years) are present in precipitation due to atmospheric testing of nuclear weapons. Most of the 25 samples shown in Fig. 3 have tritium activities that are very close to contemporary precipitation (5-7 Tritium Units) (5). However, several of the springs are relatively tritium depleted, suggesting that there is some significant mixing of recent and pre-1950 water. Hence, residence time for some of the groundwater in the basin is on the order of at least tens of years. The variation in residence time indicates that the rather high and homogeneous hydraulic conductivity indicated by the pump tests is not entirely typical of the region as a whole. However, we can generally expect rapid subsurface flow rates in the basin (velocities on the order of 100-1000 m/year).

## Stream Chemistry Relative To Historic Data

In response to the Loma Prieta earthquake, stream chemistry in the basin changed markedly (2). In July of 1992, ionic strength of the stream water near the San Lorenzo headwaters was still elevated relative to pre-earthquake data (Fig. 4) and was more sulfate rich. Given the lack of tritium depletion in the water and the high permeability of the wells in the area, it would appear likely that much of the stream water sampled in 1992 was derived from meteoric recharge that post-dated the Loma Prieta earthquake. The high level of dissolved ions in the water likely reflects long term changes in flow paths in the groundwater system that feed the streams.

In order to identify the source areas of high level dissolved ions, over 100 stream and spring samples were taken throughout the basin during the summer of 1992 (6). During this time, storm flow contributions to the streams were minor and the chemical composition of the streams could be expected to be dominated by the influx of groundwater. Dissolved ion concentrations increase from southeast to northwest (Fig. 5). This trend can be partly attributed to spatial trends in the local geology (7). In the southwest, the rocks are principally crystalline and the groundwater associated with these rocks can be expected to be low in total dissolved solids. In the northeast, the Tertiary marine shales and sandstones that are pervasive contain abundant dissolved solids. In comparison with groundwater at higher elevations, however, the stream water is typically lower in total dissolved solids and the saturation index for calcite is in the range of 0.4-0.8. The drop in saturation index indicates that calcite is being actively precipitated along the path of groundwater flow.

Superimposed upon the geologically controlled spatial trends are elevated levels of dissolved solids in discrete portions of the northeast portion of the drainage basin. These areas apparently reflect portions of the basin where the permeability changes affected by the earthquake are still present. At 11 locations, we have historical data on stream water quality taken over the time period 1973-1975 (see Fig. 1 for locations) (8). Comparison of water quality in 1992 with the highest total dissolved solids levels recorded in the historical data confirms that ionic strength in the northeast portion basin is still elevated (Fig. 6).

## Summary and Conclusions

A simple hypothesis to explain these observations is that previously lowly permeable shales in the region have undergone long term changes due to the earthquake. The increase in ionic strength in stream water would be due to the shales carrying a

significantly greater fraction of the groundwater flow to the streams. The lack of long term streamflow increases is consistent with the generally short residence times of the groundwater in the basin. The stream chemistry may be dominated by changes in groundwater flow paths, but the streamflow is governed by the magnitude of rainfall.

This bound on the time constant inferred for the permeability changes has implications for fluid flow rates in areas of active seismicity. Seismic activity, by increasing permeability locally, can allow for significant episodic fluid transport. The length of time of this permeability increase can at least be on the order of several years. Hence, faulting induced permeability increases may play a significant role in the transport of ores and hydrocarbons in the shallow crust (9). Also, if events of permeability enhancement could be identified in the rock record, they may yield valuable information on the pre-historic earthquake cycle in a region.

## References And Notes


1. J. B. Czarnecki, *EOS*, **72**, 121, 1991.

2. S. Rojstaczer and S. Wolf, S., Geology , **20**, 211 (1992); Rojstaczer and S. Wolf, *U.S. Geol. Surv. Prof. Paper,* in press.

3. Static compression induced by the earthquake would produce stream flow increases and small rises in the water level of wells (R. Muir-Wood and G. King, *J. Geophys. Res.*, in press). Upwelling of pressurized fluids from the mid-crust (R. H. Sibson, *American Geophysical Union Maurice Ewing Series,*, **4**, 593 (1981)) could not be expected to occur within 15 minutes following the earthquake and could not explain drops in the water level of wells. Compression of previously dilatant cracks (A. Nur, Geology, **2**, 217 (1974)) also cannot explain the well response.

4. The values for permeability and rates of groundwater flow are much higher than that indicated in earlier reports (J. P. Akers and L. E. Jackson, Jr., U.S. *Geol. Surv. Water Res. Invest.* **77-15**, 7 (1977). These earlier values were estimated from well log records and not measured directly.

5. The range in tritium activity in precipitation was determined from historical records from Portland, Oregon.

6. Repeat sampling was performed at 20 stations in July of 1993. While streamflow was generally 1.5 times higher, the concentration of major ionic constituents was nearly identical to the previous year. Regression analysis of the major ionic constituents on the 20 samples indicated coefficients of correlation in excess of 0.9 and regression coefficients ranging from 0.98-1.02.

7. E.E. Brabb, *U.S. Geol. Surv. Misc. Invest. Map* **I-1905**, 1 (1989).

8. Historical data were taken from M.A. Sylvester and K. J. Covey (*U.S. Geol. Surv. Water Res. Invest.* **78-19,** 61 (1978)).

9. D. Deming, *Geology*, **20**, 83 (1992); G. Garven, S. Ge, M.A. Person, and D.A. Sverjensky, *Amer. J. Sci.*, **293**, 497 (1993).

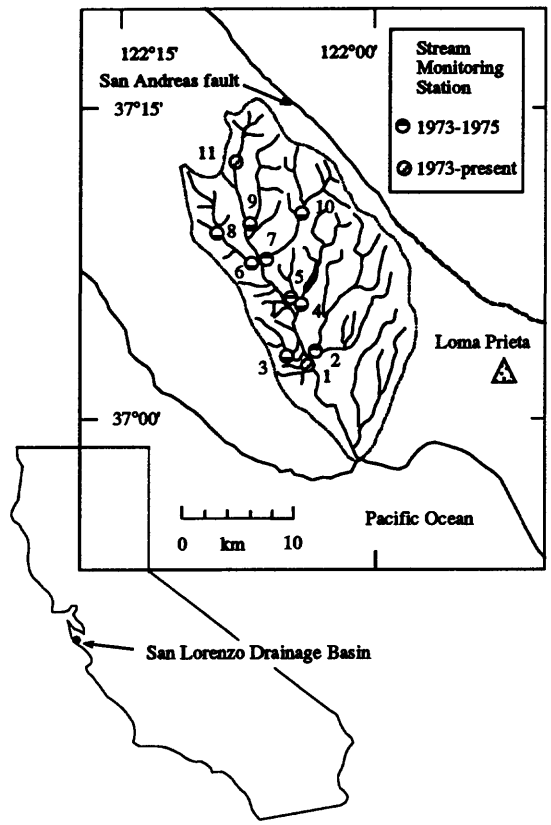



Fig. 1. Location of study area. Circles denote past or current gaging stations and the extent of water quality records at each station.

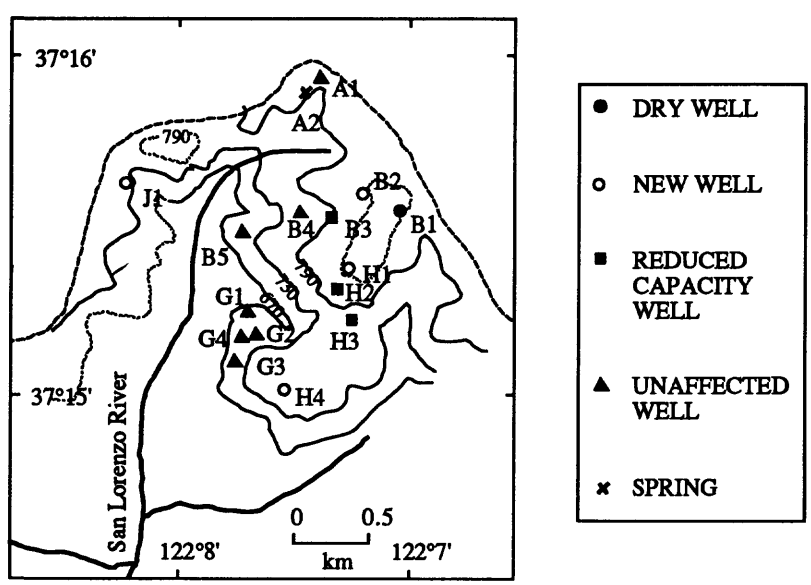



Fig. 2. Location of wells reoccupied in the summer of 1992 for hydraulic testing and water sampling.

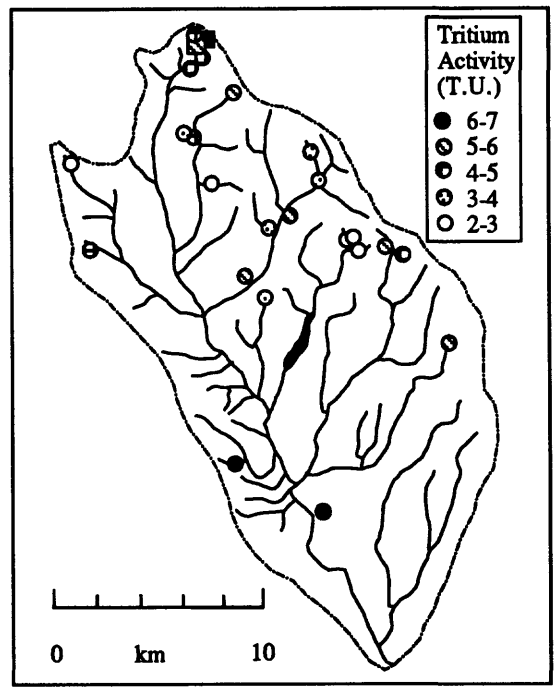



Fig. 3. Tritium activity of wells (squares) and springs (circles) in basin.

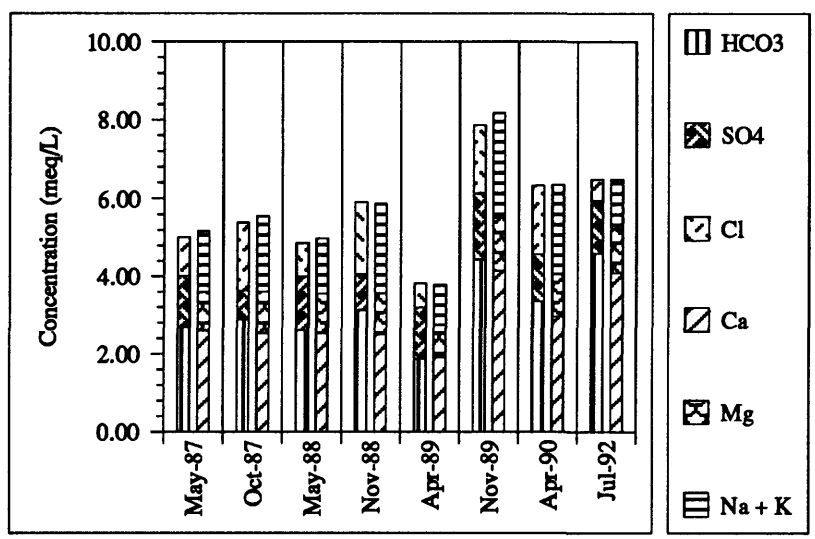



Fig. 4. Stream chemistry changes as a function of time at station 11. Location shown inFigure 1.

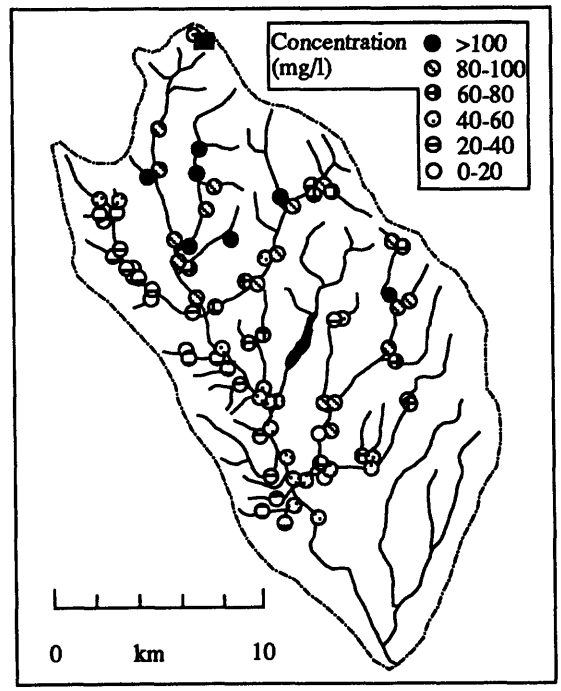



Fig. 5. Spatial trends in calcium concentration in the basin. Other major ions show similar trends.

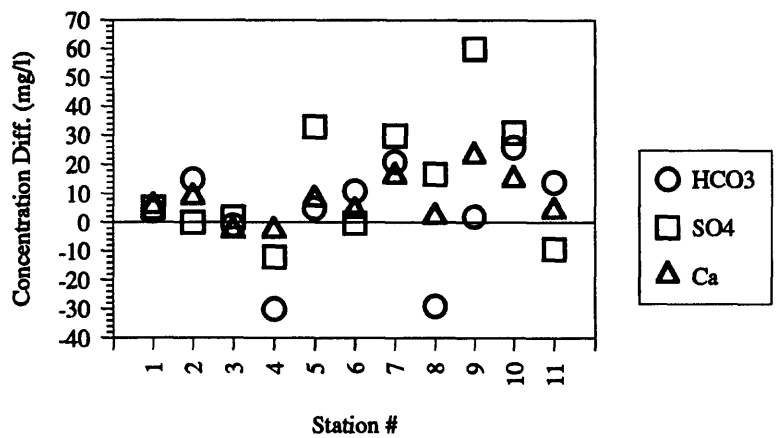



Fig. 6. Difference between dissolved solids concentrations of three major ions in 1992 and historical high levels at each station.

# 4: CHEMICAL EFFECTS OF FLUIDS ON FAULT-ZONE RHEOLOGY

# Effects of Temperature and $H_2O$ on Frictional Strength of Granite

Michael L. Blanpied, James D. Byerlee and David A. Lockner
U. S. Geological Survey, Menlo Park, CA

*Introduction*

To a first approximation, the coefficient of friction, $\mu$, of cool, faulted rocks is well represented by the combination of "Byerlee's law" [ref. 1] and the effective stress law (with the exception of certain phyllosilicate-rich aggregates). For this reason, ongoing efforts to explain the apparently low strength of some major fault zones have concentrated on ways in which fluid pressure might exceed hydrostatic in a fault at seismogenic depths. In contrast, the friction of fluid-saturated rocks at higher temperatures is not so easily represented. This is because numerous deformation mechanisms are activated or accelerated by the presence of water at elevated temperature and pressure, and several deformation processes may function simultaneously to accommodate shearing. As a result, it is difficult to predict the strength of fault zones at hydrothermal conditions, especially at depths below, say, 10 to 15 kilometers. We present here results from a suite of sliding experiments on granite "faults" designed to explore the effects of hydrothermal conditions on frictional strength and slip stability. Granite was chosen for this study because a large literature exists on its frictional properties at room temperature, as well as at elevated temperature without pore fluid pressure.

*Experimental procedure*

Velocity-stepping sliding experiments were performed in a computer-interfaced, triaxial gas apparatus on sawcut cylinders (19 mm diam.) of Westerly granite (Figure 1). The sawcut surface at 30° to the loading axis contained a layer of simulated fault gouge (granite powder ≤ 90 μm, ~0.6 mm initial thickness). Samples were jacketed with thin sleeves of annealed copper. A hole in the upper sample half provided access for pore fluid. At the start of each test, $P_c$ was raised to 400 MPa, $P_{H_2O}$ to 100 MPa, and temperature to a value between room T and 600°C. $P_p$ was held constant as axial load increased until the effective normal stress across the sliding surface, $\sigma_n^{eff}$, reached 400 MPa ($\mu$ = 0.42); thereafter $\sigma_n^{eff}$ was held constant by servo-controlled adjustment of $P_c$.

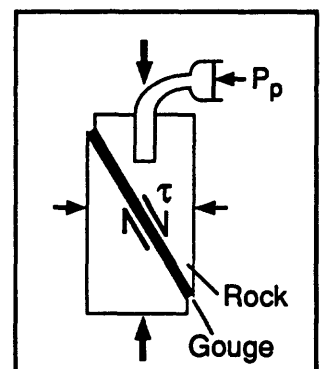


Figure 1: Geometry of sawcut granite sample.

To measure the velocity dependence of friction, the rate of piston advance was stepped up and down by a factor of ten every 0.5 mm, beginning at 2.0 mm. Velocity resolved onto the sawcut, V, was stepped between 1.0 and 0.1 μm/s, or between 0.1 and 0.01 μm/s (315 mm/year). Slip on the sawcut before jacket rupture totaled 2 to 3 mm in most tests. The experiments closely match those by Lockner et al. [2] on dried samples (theirs were run at constant $P_c$ =250 MPa, resulting in $\sigma_n$ near 400 MPa, and V was stepped between 0.05, 0.5 and 5 μm/s). For comparison of strength between tests, friction was measured at a fixed axial displacement of 2.9 mm; friction curves were interpolated in order to obtain a value at both slip rates from each experiment. Friction values presented here account for the change of contact area with slip, the pressure- and rate-dependent resistance of the O-ring seal, and the temperature-dependent strength of the copper jackets.

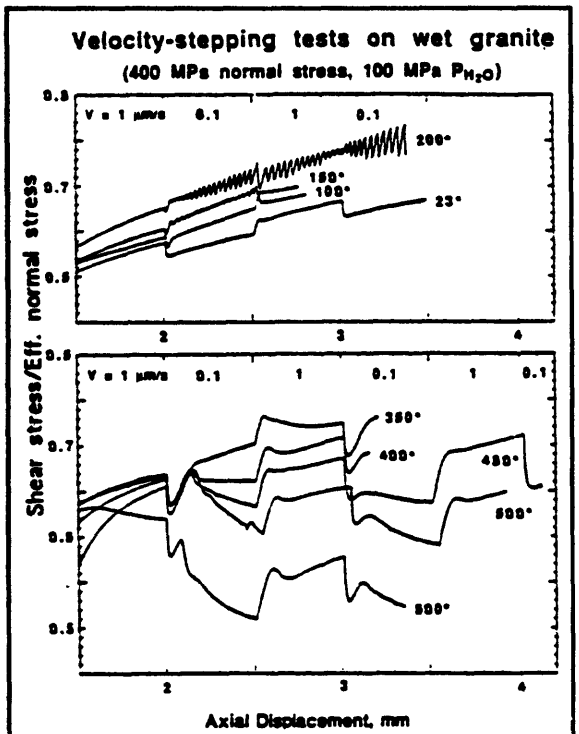



Figure 2: Friction-displacement curves for shearing of granite gouge at $P_{H_2O}$ = 100 MPa, $\sigma_n^{eff}$ = 400 MPa, and several temperatures.

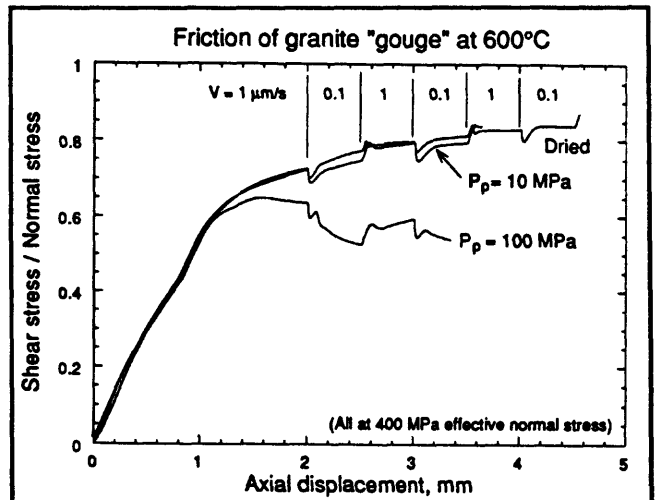



Figure 3: Friction-displacement curves for experiments at 600°C, $P_{H_2O}$ = 400 MPa, and three different pore pressures. Increasing $P_{H_2O}$ lowers the frictional strength, with a larger effect at lower slip rate.

*Results*

The strength and sliding behavior of the samples show a marked dependence on temperature, and comparison to dry tests shows that $P_{H_2O}$ affects strength above about 300°C. Representative friction-displacement curves are plotted in Figure 2. Steps in V cause a transient change in friction which decays with further slip. Below 250° samples strain harden, and marginal sliding stability is demonstrated by oscillatory slip at 150° and 200° and by small stick-slip events accompanying V steps in some tests. Similar features were noted in Lockner et al.'s dry tests. Above 250° there is a transition to strain weakening and decreasing strength with increasing temperature; strength transients are larger and more protracted. In contrast, dry tests showed continued high strength and modest strength transients to 845°.

The effect of $P_{H_2O}$ at high temperature is illustrated in Figure 3, which shows three experiments with different pore pressures but equal effective stress and temperature. The coefficient of friction in the dry run was about 0.8 and nearly insensitive to slip rate, in agreement with the results of Lockner et al. With the addition of 10 MPa $P_{H_2O}$ the friction is nearly unchanged at the higher rate, but is depressed at the lower rate. With 100 MPa $P_{H_2O}$ friction is depressed at both rates, though much more so at the lower rate, resulting in a positive rate dependence of friction. Values of friction for both wet and dry granite gouge are summarized in Figure 4. Also shown are results from Stesky [3] for the friction of fracture surfaces in dry granite at $\sigma_n^{eff}$ and V similar to ours. At low temperatures $P_{H_2O}$ has no descernable effect on friction (except via the effective stress). At high temperatures friction is inversely related to $P_{H_2O}$.

*Discussion*

The results suggest the activation of one or more water-aided deformation mechanisms at high temperature and $P_{H_2O}$, which allow slip at modest rates to proceed at friction levels well below that predicted by Byerlee's law. Deformation textures are dominated by cataclasis at the optical and SEM scales, and the search for features indicative of cooperating processes, such as solution-

transport, are hampered by the fine particle size and mix of phases. It seems possible that the weakening at hydrothermal conditions results from rapid dissolution in the presence of an undersaturated pore fluid (distilled $H_2O$): dissolution at stress-bearing particle contacts might allow shearing at stresses lower than that required for those particles to fracture. We note that the strength of the weakened samples is not a simple function of effective stress, because pore pressure plays a chemical role in determining strength as well as a mechanical one.

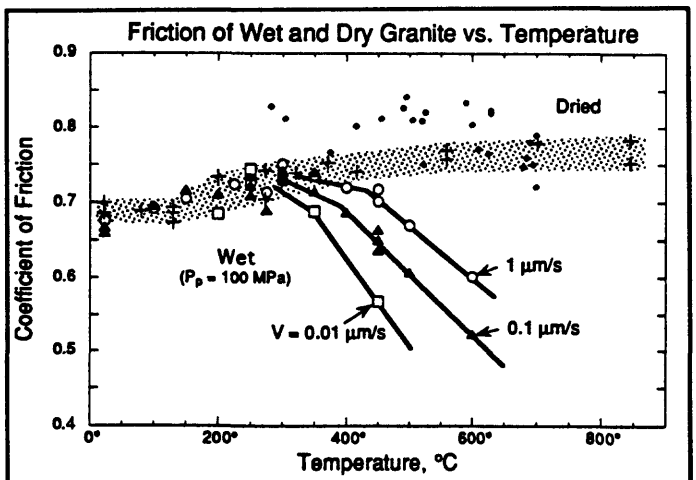



Figure 4: Friction vs. T for granite. Large symbols are friction of granite gouge with $P_{H_2O}$ = 100 MPa at three slip rates (this study). Stippled region encompasses values measured by Lockner et al. [2] for dried granite gouge. Smallest symbols are measurements for stable slip on fractures in dry granite by Stesky [3].

We have demonstrated granite friction to depend on several factors including temperature, slip rate, $P_{H_2O}$, and (to a limited extent) $\sigma_n^{eff}$ at hydrothermal conditions. Because of this complexity, and because multiple deformation processes may be active, it is difficult to extrapolate our laboratory results to conditions not encompassed by the data set. However, it is clear that Byerlee's law can be violated above 300° (depths ~9-12 km [4]), even at slip rates far higher than the San Andreas long-term rate. At these high temperatures the velocity dependence of friction is large and positive (Figure 4), whereas between ~100° and ~300° velocity dependence was generally found to be negative. In an earlier paper [5] we correlated the switch with increasing T from negative to positive velocity dependence to the transition with increasing depth from seismic to aseismic fault slip.

As a first step towards developing a constitutive description of the granite's frictional behavior we have shown previously [6] that the transient response of friction to V-steps can be modeled using the familiar rate- and state-dependent constitutive law of Ruina [7], using two state variables to represent slip-history effects. However, the values of the constitutive parameters so determined vary with conditions of temperature, $P_{H_2O}$, etc.; therefore, this approach is of limited use for extrapolation.

As a second step, we find that many features of the granite data set can be described using the rate- and temperature-dependent constitutive law of Chester and Higgs [8]. Chester and Higgs modeled data from sliding tests on ultrafine quartz powder at dry and hydrothermal conditions. They assumed that a low-T deformation mechanism (cataclasis?) and a hydrothermal mechanism (DMT?) operate simultaneously, with the latter mechanism accommodating significant strain only when temperature and $P_{H_2O}$ are high and V is low. Each mechanism is represented by a T-dependent constitutive law. Following this approach, we find that the major features of Figure 4 can be duplicated with a consistent set of constitutive parameters, although some important observations are not matched, including the gradualness of the transition between temperature weakening and temperature strengthening with increasing T, and the abrupt change from semi-stable to stable sliding at ~250°. Two further refinements in this analysis would come from: 1) performing temperature-stepping experiments [8] to constrain values of apparent activation energy, plus additional microstructural analysis to identify the important deformation mechanisms, and 2) including an explicit dependence of friction on $P_{H_2O}$ in the constitutive law for frictional slip.

Chester [this volume] has extended the analysis of quartz friction further, separating the low-temperature regime into two regimes with contrasting velocity- and temperature dependence. Although following this step would no doubt improve the fit to the granite data, the resulting proliferation of parameters is not justified in this stage of the granite analysis. It is conceivable that the two- (or three-) mechanism analysis, which appears to work well for quartz, may not turn out to be appropriate for granite because of its mixture of mineral phases: Feldspar, for example, may define two regimes of frictional deformation with a boundary distinct from that defined by quartz. This would presumably lead to a broadening of the transition region, not easily represented by Chester's formulation.

## References:


[1] Byerlee, J., Friction of rocks, *Pure Appl. Geophys.*, *116*, 615-626, 1978.

[2] Lockner, D. A., R. Summers and J. D. Byerlee, Laboratory measurements of velocity-dependent frictional strength, *Alternate Journal*, 1986.

[3] Stesky, R. M., W. F. Brace, D. K. Riley and P.-Y. F. Robin, Friction in faulted rock at high temperature and pressure, *Tectonophysics*, *23*, 177-203, 1974.

[4] C. Williams, A. Lachenbruch and J. Sass, unpublished geotherms for the California Coast Ranges, 1992.

[5] Blanpied, M. L., D. A. Lockner and J. D. Byerlee, Fault stability inferred from granite sliding experiments at hydrothermal conditions, *Geophys. Res. Lett.*, *18*, 609-612, 1991.

[6] Lockner, D. A., M. L. Blanpied, C. Marone and J. D. Byerlee, Friction constitutive parameters for sliding of granite at hydrothermal conditions (abstract), *EOS, Trans AGU*, *72*, 456, 1991.

[7] Ruina, A. L., Slip instability and state variable friction laws, *J. Geophys. Res.*, *88*, 10359-10370, 1983.

[8] Chester, F. M. and N. G. Higgs, Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions, *J. Geophys. Res.*, *97*, 1859-1870, 1992.

# Rheologic Model of Crustal Faults: Influence of Fluids on Strength and Stability

FREDERICK M. CHESTER

*(Department of Earth & Atmospheric Sciences, 3507 Laclede Ave., Saint Louis University, St. Louis MO 63103)*

## ABSTRACT

Current two-layer rheologic descriptions of the crust are inadequate in the brittle to ductile transition regime because they neglect mixed-mechanism modes of deformation and solution-transfer assisted processes. On the basis of rock deformation experiments on wet quartz, a rheologic model incorporating multiple frictional slip mechanisms and dislocation creep is developed to address strength and stability of faulting in the crust. The model incorporates rate and temperature dependence of friction for several slip mechanisms including localized slip and solution-transfer assisted slip. The solution transfer mechanism can lead to a significant reduction of strength where it dominates in the brittle to ductile transitional regime in the mid-crust. Rate weakening friction behavior necessary for the nucleation of seismic slip is associated with the localized slip mechanism that dominates at intermediate temperature and slip rates. Application of the rheologic model to the San Andreas fault indicates the transition from rate weakening to rate strengthening occurs within the frictional regime and is both rate and temperature dependent. Depending on fault thickness the transition can occur at depths to 12 km approximately consistent with the depth of seismogenesis on the San Andreas fault. The basic framework for treating multiple slip mechanisms in rheologic models will be important as our understanding of mixed-mechanism behavior improves.

## INTRODUCTION

Changes in pressure, temperature, rock type, and pore-fluid conditions with depth in the crust and mantle produce rheologic layering in the outermost shell of the earth. Geologic studies have long indicated that there is a change from brittle to ductile deformation behavior with depth in the Earth's crust. This change in deformation style is explained with a two-layer rheologic model of the crust based on the results of experimental rock deformation studies in the late 1970's [Goetze & Evans, 1979; Brace & Kohlstedt, 1980]. The rheologic model is based on two important findings: 1) the frictional strength of fractures and faults in rock is largely independent of all parameters except pressure and can be described adequately using a simple Coulomb relation (e.g., Byerlee's Law)[Byerlee, 1978]; 2) rock can deform ductily at high temperature and low strain rates by intracrystalline plastic mechanisms (e.g., dislocation creep) that give rise to a non-linear viscous behavior that is largely pressure independent. Thus, the shallow brittle crust will display frictional behavior at low pressures and temperatures (Figure 1). Strength increases with depth as pressure increases until dislocation creep is activated at the elevated temperatures at mid-crustal depths. In the simplest case this model correlates the transition from friction to dislocation creep behavior to strong crust. The two-layer rheologic model has been used over the past 15 years with little modification to understand crustal faulting, lithospheric flexure, mountain building and other geodynamic phenomena [e.g., Goetze & Evans, 1979; Sibson, 1983; Davis et al., 1983].

In detail, there are a few important aspects of rheology that the two-layer model does not address. Most importantly, the model does not consider several deformation mechanisms that may be important in the crust and accordingly overestimates the strength of the crust [e.g., Sibson, 1984]. In addition, the model does not describe the transient behavior that is relevant to fault stability, i.e. seismogenesis [Scholz, 1990]. As a result the model may be too simplistic and potentially misleading. In many cases, incorrect associations and conclusions are made in conjunction with using the model [e.g., Rutter, 1986]. These include, equating frictional behavior with brittle behavior, equating brittle (or frictional) behavior with seismogenic behavior, assuming

that the brittle to ductile transition is abrupt, equating seismicity with high strength, and assuming that the crust at the transition is very strong (Figure 1).

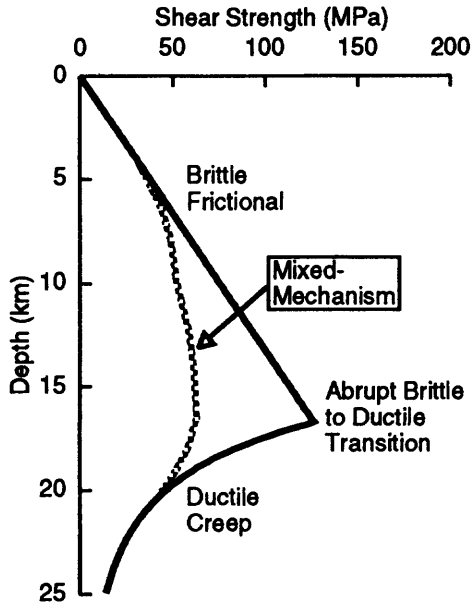



Figure 1. An example of a shear strength profile (solid line) for the crust based on a two-layer rheologic model. The effect of incorporating a mixed mechanism rheology on strength of the mid-crust is shown schematically with the broken line.

The change from brittle friction to dislocation creep in the crust is often referred to as the brittle-ductile transition (Figure 1). Many studies indicate that this transition occurs over a broad depth range in the crust and is characterized by mixed-mechanism deformation [Kirby, 1983; Chester, 1988; Scholz, 1990]. Chester [1989] has shown that within the transitional regime, brittle and intracrystalline plastic mechanisms can be coupled to produce a lower strength than if both mechanisms were to operate independently. In addition, the operation of fluid-assisted diffusive mass transfer processes within the transitional regime should lead to further reductions in crustal strength [e.g., Sibson, 1984; Chester & Higgs, 1992]. A complete rheologic description of mixed-mechanism deformation for the transitional regime is not available, but we do expect that it will display pressure-, temperature- and rate-dependent strength. Although the pressure-dependent strength will constitute a frictional behavior, the mixed-mechanism deformation will not be the same as that occurring in the shallow crust, and therefor probably will not follow Byerlee's Law.

Since Brace and Byerlee (1966) presented the stick-slip hypothesis for earthquakes there has been extensive research on the stability of frictional systems. It is now recognized that the stability of a fault is independent of the total strength of the fault. Apparently, stability depends on a small magnitude slip-rate dependence of strength [e.g., Tse & Rice, 1986]. Rate-weakening behavior, i.e., a decrease in frictional strength with an increase in slip rate, is generally necessary for seismic slip [Dieterich, 1978; Gu et al., 1984]. Rock friction experimental results have led to the suggestion that stability transitions in faults, such as at the base of the seismogenic layer of the crust, reflect small changes in the rate dependence of frictional strength and occur entirely within the frictional regime. These data suggest that the correlation of the base of the seismogenic zone with the transition from friction to dislocation creep may be incorrect [Tse & Rice, 1986; Chester, 1988; Scholz, 1990].

The purpose of this paper is to extend rheologic models for the crust through a multi-mechanism model of friction as expressed in the deformation-mechanism map format [e.g., Frost & Ashby, 1982; Chester, 1988; Chester & Higgs, 1992]. This treatment illustrates a basic framework for incorporating multiple friction and creep mechanisms and for characterizing behavior of crustal faults at all depths over the earthquake cycle. The most important aspect is the addition of a solution-transfer assisted frictional-slip mechanism that may be appropriate for describing the brittle to ductile transitional regime at mid-crustal depths. In addition, the rate dependence of the friction mechanisms is described so that questions of fault stability may be

addressed. The model is applied to the San Andreas fault to investigate the strength and stability of crustal faults.

## MECHANISMS OF FRICTION

We assume that faulting in the crust occurs by several distinct slip mechanisms that contribute to the total slip, but at most conditions of pressure, temperature and slip rate, one of the slip mechanisms will dominate and therefore govern the gross mechanical behavior of the fault. This approach is motivated by experimental observations that illustrate that distinctly different mechanical behavior and associated fault gouge fabrics are produced at different environmental conditions of shearing [e.g., Chester & Logan, 1990; Chester & Higgs, 1992].

Although an understanding of the exact mechanisms that produce rate, temperature and slip-history dependence of frictional strength is lacking, there is sound evidence that different frictional slip mechanisms can operate within fault gouge layers. Many experimental studies have demonstrated that the rate-dependent friction properties of granular material change with shear strain [e.g., Dieterich, 1981; Marone & Kilgore, 1993]. This change correlates with the progressive slip localization within shearing granular layers. In experimental configurations that promote distributed shear and comminution within gouge layers, for example by using coarse grained porous aggregates or rough bounding surfaces, the gouge displays rate strengthening behavior. Microstructural studies of quartz and granite gouge sheared at these conditions demonstrate that extensive cataclastic grain size reduction occurred [e.g., Marone and Scholz, 1989; Biegel et al., 1989] by constrained comminution [Sammis et al., 1987]. In experimental configurations where localized slip is favored, for example by shearing layers bounded by smooth surfaces or after shearing to large displacements and causing extensive grain size reduction, rate weakening behavior is displayed. In granular halite gouge the transition from distributed shear and rate strengthening behavior to localized slip and rate weakening behavior is not only dependent on shear strain, but also on slip-rate and normal stress [Chester and Logan, 1990]. These observations are used to justify treating the distributed shearing and localized slip as distinct frictional slip mechanisms. For descriptive purposes, the mechanism typical of experiments at low shear strain characterized by distributed grain size reduction and rate-strengthening behavior is referred to as the cataclastic flow (CF) mechanism. The mechanism characterized by slip localization and rate-weakening behavior is referred to as the localized slip (LS) mechanism.

Excellent evidence for an additional frictional slip mechanism comes from experiments on faulted rock and gouge layers at elevated temperatures. Several experimental studies have demonstrated pronounced weakening for frictional slip at low rates and elevated temperatures in the presence of pore water (Figure 2)[e.g., Rutter and Mainprice, 1978; Higgs, 1981; Dula, 1985]. Recent rate stepping experiments on granite gouge under dry [Lockner et al., 1986] and water saturated [Blanpied et al., 1991] conditions also demonstrate pronounced weakening by water at elevated temperatures (Figure 3). In shear experiments on ultrafine quartz gouge, Higgs [1981] showed that the onset of water weakening coincides with a major change in the fabric of the gouge. Gouge sheared in the water weakened regime displays distributed deformation, generation of preferred crystallographic fabrics and vein formation, all of which are taken as evidence of solution-transfer processes. At lower temperatures where water weakening was not observed the deformation was localized to discrete slip surfaces consistent with the LS mechanism. Thus, microstructural observations indicate that the weakening effect of water is due to the activation of a new slip mechanism involving solution-transfer processes [Rutter and White, 1979; Higgs, 1981]. For descriptive purposes, this slip mechanism is referred to as the solution-transfer (ST) mechanism.

Early efforts to model the water weakening effect utilized creep laws for a diffusive mass transfer (pressure solution) slip process [Rutter and Mainprice, 1979]. However, several observations indicate that the water-weakening occurs entirely within the frictional regime and should be described by friction constitutive relations. A dependence of shear strength on effective pressure (i.e., frictional behavior) for the ST mechanism was clearly demonstrated in high temperature shear experiments on quartz gouge by Dula [1985] (Figure 2). Moreover, slip rate stepping experiments by both Dula [1985] and Blanpied et al [1991] show that changes in slip rate

produce transient behavior typical of rock friction. Chester & Higgs [1992] demonstrated that the transient strength response of quartz gouge to changes in slip rate for both the LS and ST mechanisms, as well as across the transition between mechanisms can be described using friction constitutive relations.

## MULTIMECHANISM FRICTION CONSTITUTIVE MODEL

We assume all friction mechanisms can be described by the same constitutive relation, but each mechanism is characterized by different constitutive parameters. The friction mechanisms are described with the state variable friction constitutive relation developed by Dieterich [1979] and Ruina [1983] but modified to include temperature dependence as described by Chester [1994]. The constitutive relation is given by

$$\mu = \mu^* + A\left[\ln\left(\frac{V}{V^*}\right) + \frac{Q_A}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)\right] + B\Theta$$

$$\frac{d\Theta}{dt} = -\frac{V}{D_c}\left[\Theta + \ln\left(\frac{V}{V^*}\right) + \frac{Q_B}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)\right],$$

where $\Theta$ is the state variable, $\mu$ is the coefficient of friction, $V$ is slip velocity, and $T$ is temperature. The constitutive parameter $A$ describes the magnitude of the direct dependence of friction on slip rate. Parameters $B$ and $D_c$ describe the magnitude and characteristic slip distance for the the slip-history dependence. Parameters $Q_A$ and $Q_B$ are the apparent activation energies for the mechanisms controlling rate dependence. Chester [1994] shows that the constitutive relation can adequately describe the transient frictional behavior of quartz gouge during shear following sudden changes in temperature and slip rate (Figure 4).

At steady-state the constitutive relation has the form

$$\mu^{ss} = \mu^* + (A - B)\ln\left(\frac{V}{V^*}\right) + \frac{(AQ_A - BQ_B)}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)$$

where rate-dependence is

$$\frac{d\mu^{ss}}{d(\ln V)} = A - B$$

temperature-dependence is

$$\frac{d\mu^{ss}}{d(1/T)} = (AQ_A - BQ_B)/R$$

and $\mu^{ss} = \mu^*$ at $V = V^*$ and $T = T^*$.

At this time the data on the rate and temperature dependence of friction is limited to quartz gouge [Chester & Higgs, 1992; Chester, 1994] and wet granite gouge [Blanpied et al., 1991, this volume]. The constitutive parameters for frictional behavior of wet quartz gouge in the three mechanism fields are listed in Table 1.

Chester [1993] conducted rate and temperature stepping experiments on wet granular quartz layers sheared between rough surfaces at temperatures to 82 °C. The experiments display rate strengthening behavior consistent with earlier experiments on wet quartz at room temperature by Marone et al. [1990] (Figure 4). In addition, the apparent activation energies for the friction mechanisms, $Q_A$ and $Q_B$, are approximately the same magnitude, 89 kJ/mol. Thus, as would be expected based on the equations above, the quartz gouge is temperature-weakening (Figure 4). On the basis of the microstructures of wet quartz gouge documented by Marone & Scholz [1989], the

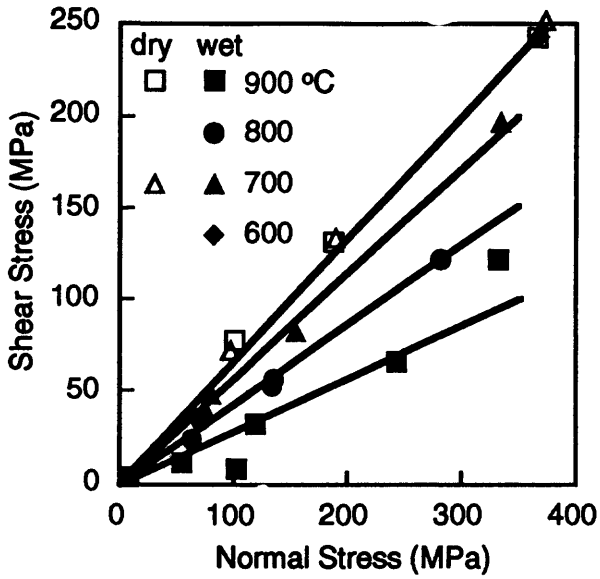



Figure 2. Shear stress versus effective normal stress for wet and dry quartz gouge at a shear strain rate of $10^{-3}$/s and temperatures between 600 and 900 °C as reported by Dula [1985]. Approximately linear dependence of shear strength on normal stress shows frictional behavior at all conditions. Comparison of strengths for wet and dry conditions displays the weakening effect of water at temperatures above 600 °C.

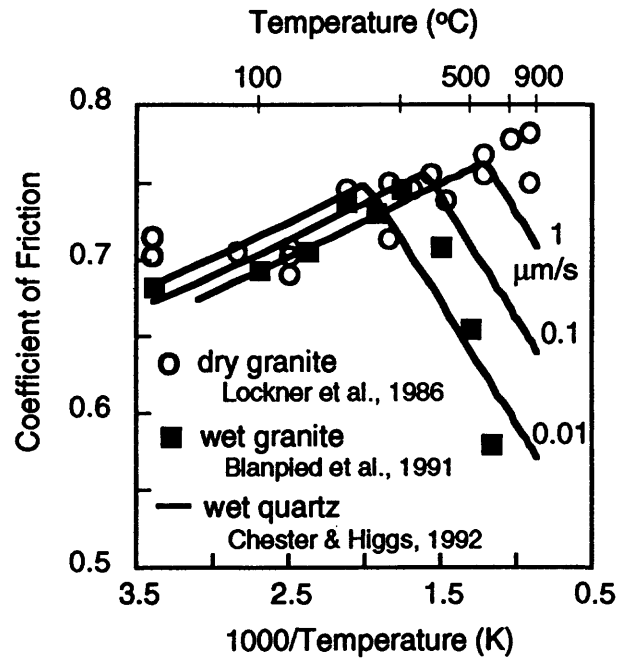



Figure 3. Dependence of friction on temperature for dry granite gouge [Lockner et al., 1986], wet granite gouge [Blanpied et al., 1991] and wet quartz gouge [Chester and Higgs, 1992]. Comparison of wet and dry granite gouge friction shows water-weakening at temperatures above approximately 300 °C. Friction of wet quartz gouge shown for three different rates of slip (1, 0.1 and 0.01 µm/s). Wet granite and quartz gouge show similar temperature dependencies at low and high temperature conditions.

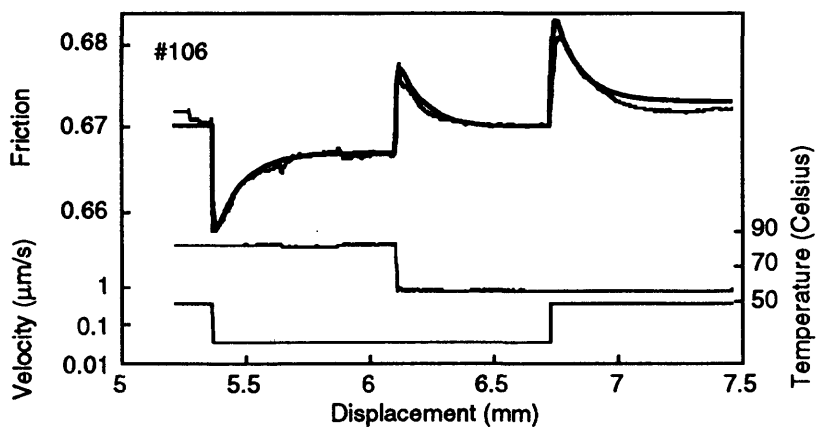



Figure 4. Example of the results and model simulations of slip velocity and temperature stepping experiments on wet quartz gouge [Chester, 1994]. Sequence of velocity and temperature steps imposed on the gouge layers are shown. The friction response of quartz gouge is indicated by the thin lines (observed) and thick lines (model). The model is based on the state variable friction constitutive relation described in the text.

rate and temperature dependent frictional behavior observed by Chester [1994] is interpreted to represent the CF mechanism.

Table 1. Constitutive Parameters for Friction in Wet Quartz Gouge

| Mechanism | $A$ | $B$ | $D_C$ (μm) | $Q_A$ (kJ/mol) | $Q_B$ (kJ/mol) | $\mu^*$ | $\frac{d\mu^{ss}}{d(\ln V)}$ | $\frac{d\mu^{ss}}{d(1/T)}$ | $V^*$ |
|---|---|---|---|---|---|---|---|---|---|
| CF | 0.0062 | 0.0042 | 130 | 89 | 89 | 0.638 | 0.002 | 21.4 | $t(10^{-3})$ |
| LS | 0.006 | 0.011 | 20 | 78 (?) | 78 (?) | 0.737 | -0.005 | -47.4 | 1 |
| ST | 0.03 | 0 | - | 44 | - | 0.850 | 0.030 | 159 | $t(10^{-3})$ |

$V$ and $T$ are in units of μm/s and Kelvin, respectively, $t$ is gouge layer thickness in units of μm, and $T^* = 573$ K.

Chester and Higgs [1992] analyzed the transient mechanical response of wet quartz gouge during slide-hold-slide tests at temperatures to 600 °C. As described above, these experiments display microstructures and mechanical behavior across the transition from the LS to ST slip mechanism. At the higher temperature and lower slip rate conditions where the ST mechanism dominates, the gouge displayed rate strengthening and temperature weakening behavior (Figure 3). There is little slip history dependence ($B = 0$) for this mechanism, and the apparent activation energy $Q_A$ is approximately 44 kJ/mol consistent with solution transfer as rate controlling process [Chester & Higgs, 1992]. At the lower temperature and higher slip rate conditions, the gouge displays rate weakening and temperature strengthening behavior (Figure 3). The experiments were not designed to determine values of $Q_A$ and $Q_B$. However assuming $Q_A$ and $Q_B$ are equivalent, as found for the CF mechanism by Chester [1994], the rate dependence and steady-state temperature dependence displayed in the experiments implies an activation energy of approximately 78 kJ/mol. Thus, the activation energy for both the CF and LS mechanisms may be approximately the same.

Together, the results of experiments on wet quartz gouge indicate that rate weakening behavior occurs only for intermediate conditions of slip rate and temperature (Figure 5). At very large or very small temperatures or slip rates, rate strengthening behavior is observed either due to the operation of the CF or ST mechanisms. The transition from rate weakening to rate strengthening depends on both temperature and slip rate (Figure 6). The experiments on wet granite gouge by Blanpied et al. [1991, this volume] document a transition at elevated temperatures from rate-weakening and temperature-strengthening to rate-strengthening and temperature-weakening, similar to the transition seen in the quartz between the LS and the ST mechanisms. In fact the magnitude of the rate and temperature dependencies are roughly equivalent for granite and quartz in both mechanism fields (Figure 3). The most notable difference is that the transition between the LS and ST mechanisms in the granite gouge occurs at a slightly lower temperature than for the wet granite gouge (for the same gouge thickness and slip rate).

Assuming we have successfully identified and characterized the important friction mechanisms, there remains at least two general questions that must be addressed when constructing a rheologic model for the crust: 1) Must the friction parameters be scaled for application to natural deformation? 2) How do the individual mechanisms operate in combination?

It has long been suspected that the critical slip distance, $D_C$, should scale with some geometric characteristic of faults because the parameter has units of length. Recently Marone & Kilgore [1993] present evidence that $D_C$ represents a characteristic strain within the actively shearing zone of material. Thus, $D_C$ should scale with active shear band thickness. Whether rate dependent parameters $A$ and $B$ scale with thickness is not yet known. A related question is whether friction constitutive relations should be cast in terms of slip rate or shear strain rate. In the case of the ST

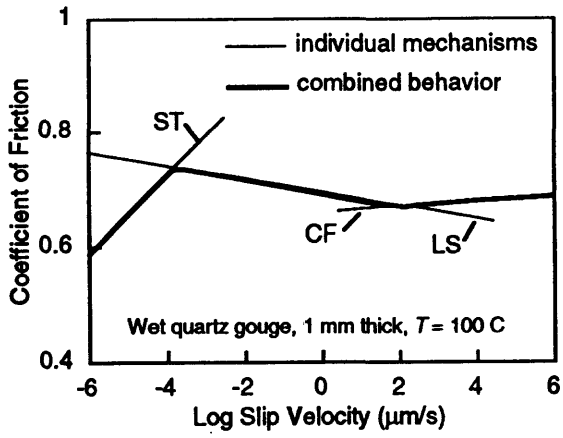



Figure 5. Friction constitutive behavior of the individual slip mechanisms in wet quartz gouge as observed in experiments and described by the friction constitutive relations in Table 1. The ST and CF mechanisms are rate-strengthening. Rate weakening behavior due to the LS mechanism occurs at intermediate slip rates. The combined behavior for slip in gouge by all three mechanisms as described by the multimechanism friction constitutive model is represented with the heavy line.

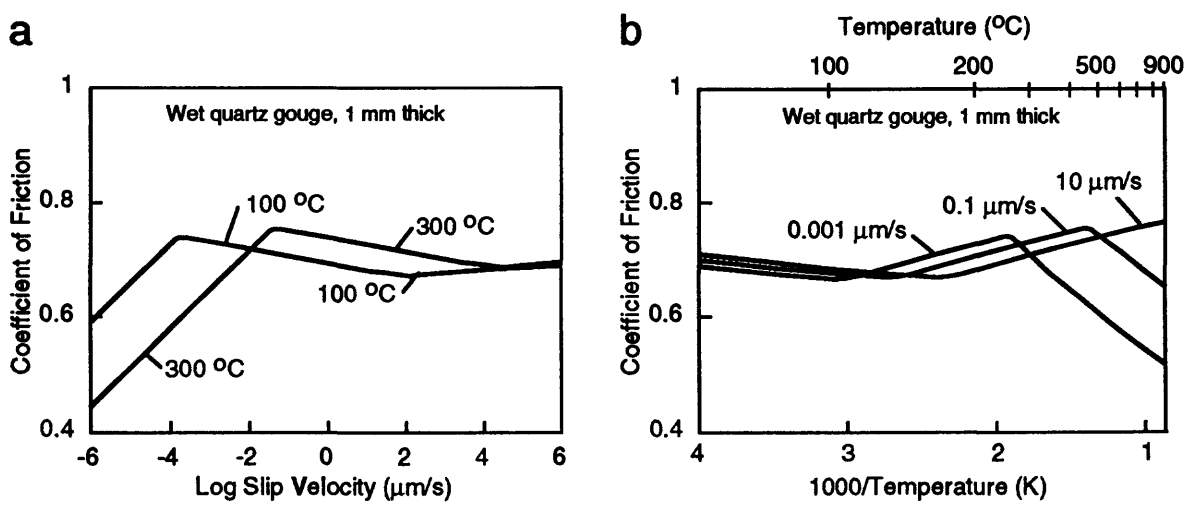



Figure 6. Friction of wet quartz gouge described by the multimechanism model as a function of temperature and slip rate. Transitions in rate dependence depend are functions of rate and temperature.

mechanism, deformation in the gouge layer is distributed in a homogeneous fashion [Higgs, 1981]. Thus, it would appear most appropriate to express the constitutive law in terms of shear strain rate and take into account the thickness of the gouge layer. In contrast, the LS mechanism appears to result in all slip being localized to a discrete surface or set of surfaces. In this latter case the thickness of the gouge layer is not important and the friction law should be cast in terms of slip rate. There is some logic to assuming distributed deformation for rate-strengthening behavior and localized slip for rate-weakening mechanisms, at least at steady-state. However, all of the questions raised herein concern aspects of gouge deformation that are not entirely understood. For the purposes of this analysis, it is assumed that the constitutive relations for mechanisms having rate-strengthening behavior should take into account the thickness of the deforming layer, whereas rate-weakening relations should not include gouge thickness. This is done simply by including gouge thickness as a multiplicative factor for the reference velocity, $V^*$ (Table 1). Other possible scaling effects are not treated at this time.

The exact manner that the constitutive relations for the mechanisms should be combined to describe frictional slip is unknown because there have not been sufficient experiments conducted in the transitional regimes and because the individual mechanisms are not completely understood. In general, the relations for each mechanism are combined assuming that the mechanisms operate in either a parallel-concurrent or series-sequential fashion [Poirier, 1985]. One assumption is that the CF and the LS mechanisms operate series-sequential. This is implied by the observation that at any particular conditions of slip rate and temperature, the mechanism giving rise to the greatest coefficient of friction is the dominant mechanism (Figure 5). Series-sequential behavior implies that both mechanisms must operate in order for slip to occur. Sammis [1993], Marone et al. [1990] and others have modeled rate-dependent friction in terms of distinct microscopic processes involving particle fracture, interparticle slip, and dilation. These models are consistent with the view taken here in which constrained comminution (involving fracture) is distinct from localized (interparticle) slip but that both must operate to some extent for slip to occur during brittle deformation of a gouge layer. A second assumption is that the ST mechanism operates parallel concurrent with the CF and LS mechanisms because at any particular conditions of rate and temperature the weaker mechanism dominates (Figure 5).

The customary assumptions that strain (or slip) rates sum for parallel concurrent mechanisms and the reciprocal of strain rates sum for series sequential mechanisms [e.g., Poirier, 1985] do not strictly apply for the case described herein where both rate strengthening and rate weakening mechanisms are involved. We make the simplest assumption that the overall frictional behavior of gouge is equivalent to the behavior of the dominant mechanism. The ST mechanism is dominant if $\mu_{ST}(T,V) < \mu_{LS}(T,V)$, the CF mechanism is dominant if $\mu_{CF}(T,V) > \mu_{LS}(T,V)$, and the LS mechanism is dominant if $\mu_{CF}(T,V) \leq \mu_{LS}(T,V) \leq \mu_{ST}(T,V)$. The boundaries between dominant mechanism fields, that is the rate and temperature conditions at which the dominant mechanism changes, are defined as the temperature and rate where two mechanisms produce equivalent coefficients of friction.

## RHEOLOGIC MODEL FOR CRUSTAL FAULTING

The rheologic model is generated by combining the multimechanism friction constitutive relation with a relation for viscous flow by dislocation creep. The friction and creep mechanisms are treated as operating parallel-concurrent so that the mechanism giving rise to the lowest shear stress will be dominant. The model uses the power-law creep relation for wet quartz as determined by Koch et al. [1989] but it has been rewritten for simple shear to be appropriate for faulting. The creep relation is

$$\dot{\gamma} = 3^{(n+1)/2} C \, \tau^n exp[-H/RT]$$

where $\dot{\gamma}$ is shear strain rate, $\tau$ is shear stress, $C = 5.05 \times 10^{-6}$ MPa$^{-n}$ s$^{-1}$, $H = 145$ kJ/mol, and $n = 2.61$.

The shear strength of a frictional fault depends on the type of fault, effective pressure, and the coefficient of friction [Sibson, 1974]. For the case of a strike-slip fault, the minimum differential stress to activate slip is given by

$$(\sigma_1 - \sigma_3) = 2\sigma_v^e(R' - 1)/(R' + 1)$$

and

$$R' = [-\mu + (1 + \mu^2)^{1/2}]^{-2}$$

assuming that the effective overburden pressure ($\sigma_v^e$) is equal to the intermediate principal stress and to the mean stress, i.e.,

$$\sigma_v^e = \sigma_2^e = (\sigma_1^e + \sigma_3^e)/2$$

and

$$\sigma_v^e = \rho g z(1-\lambda),$$

where $\lambda$ is the pore fluid factor defined as the ratio of pore fluid pressure to overburden pressure,

$$\lambda = P_f/\rho g z.$$

In this case the minimum shear stress to activate frictional sliding is

$$\tau = \{(\sigma_1 - \sigma_3) \sin [\tan^{-1}(1/\mu)]\}/2.$$

The implications of the rheologic model to the strength and stability of crustal faults is illustrated by considering a strike-slip fault with slip rate and temperature conditions appropriate to the San Andreas fault. Parameters assumed for the model are a geotherm of 20 °C/km, overburden pressure gradient of 25 MPa/km, and an average slip rate of 30 mm/yr. The overall thickness of the fault zone in the creep regime is fixed at 100 m, however, within the frictional regime we consider both thin (1 mm) and thick (10 m) fault gouge zones.

The shear strength of the strike-slip fault is shown as a function of depth and pore fluid factor for cases of thin and thick gouge layers and for sliding at a constant slip rate of 30 mm/yr (Figure 7). The plots illustrate the role of pore fluid pressure in reducing fault strength in the frictional regime. In addition, an increase in gouge layer thickness weakens a fault in the frictional regime due to the strain-rate dependence of the ST mechanism. A fault with a 10 m thick gouge layer is significantly weaker than that often assumed based on Byerlee's Law (Figure 8). We suggest that the rheologic model herein represents an upper bound on fault strength because it is based on a monomineralic quartz rheology. Comparison of quartz gouge and granite gouge friction suggests that the ST mechanism should be dominant in polymineralic gouge at lower temperatures and higher rates relative to quartz gouge, and therefore should be weaker overall. Moreover, operation of crystal plasticity in conjunction with solution transfer and cataclasis would probably further reduce frictional strength in the transitional regime. Application of the rheologic model to the San Andreas fault suggests modest increases in pore fluid factor are necessary to reduce the shear strength to levels required by heat-flow and directional constraints [e.g., Lachenbruch and McGarr, 1990].

The rheologic aspects of fault stability are investigated by delineating the rate and depth conditions over which the different mechanisms dominate in the model fault (Figure 9). At a fixed

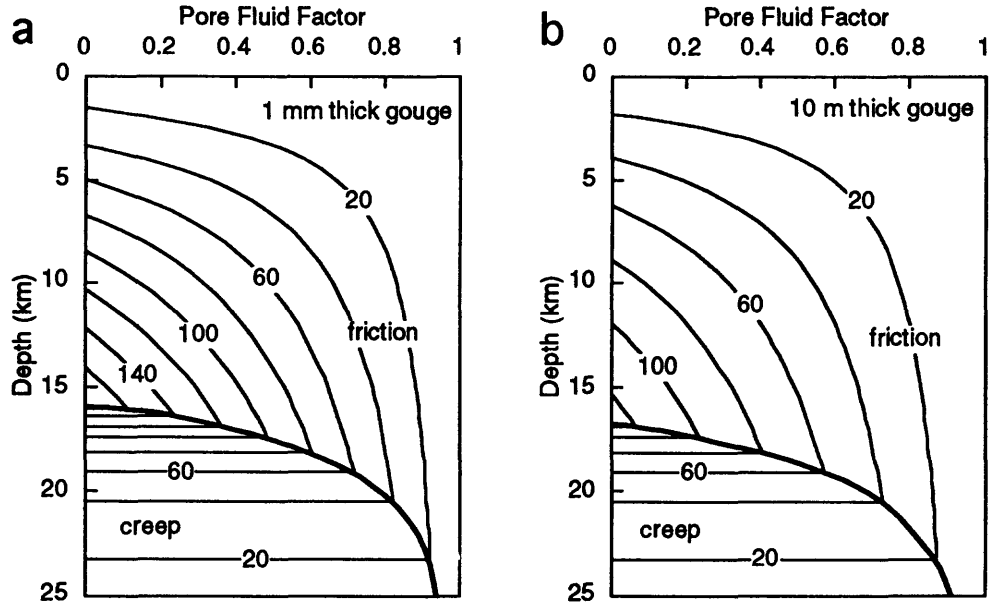



Figure 7. Shear strength (MPa) of the San Andreas fault as a function of depth and pore fluid factor for (a) 1-mm thick and (b) 10-m thick gouge zones based on the rheologic model for wet quartz. See text for further description.

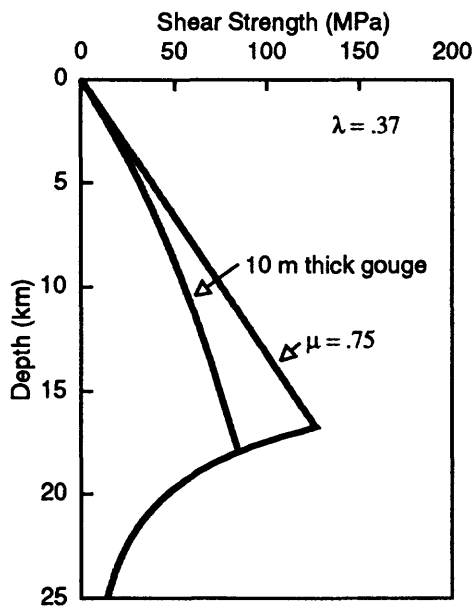



Figure 8. Shear strength as a function of depth for the rheologic model of the San Andreas fault with a 10 m thick gouge zone and for a simple two-layer rheologic model using a constant coefficient of friction of 0.75 for the frictional regime. The weakening produced by the addition of the ST mechanism within a thick gouge zone is apparent.

depth, faults will display rate strengthening behavior at very low slip rates due to the ST mechanism. Higher slip rates will allow the LS mechanism to dominate and change behavior to rate weakening. At the highest slip rates and sufficiently shallow depths, the fault may revert back to rate strengthening behavior from the CF mechanism. As mentioned previously, many experiments indicate that the transition from the LS to CF mechanisms is strain dependent, with the LS mechanism field expanding at the expense of the CF mechanism field with greater strain. This suggests that for mature faults the CF mechanism may not be significant except at very high slip rates and shallow depths.

The rate weakening behavior of the LS mechanism is a necessary condition for nucleation and propagation of seismic slip. The rate strengthening behavior of the ST mechanism should produce stability. At the average slip rates of the San Andreas, the ST mechanism is dominant throughout most of the frictional regime for faults with thick gouge layers (Figure 9). Thus, thicker faults should tend to display aseismic creeping behavior throughout the crust except at very shallow depths. The conclusion that faults with thick gouge layers will tend to be inherently stable also was reached by Marone and Kilgore [1993] based on the scaling of the critical slip distance, $D_C$, on gouge thickness and the influence of $D_C$ on stability. In our analysis the influence of gouge thickness on stability arises from the fact that strain rates decrease with increasing thickness, and lower strain rates promote the weaker, more stable ST friction mechanism.

At the average slip rate of the San Andreas fault the LS mechanism is dominant at upper crustal levels for faults with thin gouge layers (Figure 9). Extremely thin gouge layers or extreme localization of slip to discrete surfaces within a gouge layer are required in order for the rheologic model to predict rate weakening behavior over the entire seismogenic regime of the San Andreas fault (to depths of approximately 12 km). The gouge thickness specified herein does not refer to the entire thickness of the fault, only the zone that is actively shearing at any particular time. Observations of ancient faults clearly indicate the presence of discrete slip surfaces within thicker gouge layers [Chester et al., 1993; Chester, 1993]. It is reasonable to assume that shear displacements are extremely localized during seismic slip and that the active shearing zone may be very thin throughout the seismic regime. Such localization could explain the nucleation of seismic slip within the San Andreas near the base of the seismogenic regime using this rheologic model.

The propensity for seismic slip as a function of depth also is qualitatively indicated by comparing the strength of the modeled fault at the average slip rate for the San Andreas with the strength at coseismic slip rates (Figure 10). For 1 mm thick gouge layers the model predicts a stress drop will occur at depths up to 12 km, which is consistent with the inferred depth of earthquake rupture along much of the San Andreas fault. Recognizing the possible dependence of the LS to CF transition on strain, the magnitude of stress drops and the depth range of stress drops should be increased for models of mature faults in which the CF mechanism does not operate. For thick faults (10 m) the model predicts that stresses would increase during seismic slip suggesting again that increasing fault thickness promotes stability. The analysis of stability using this rheologic model only considers the rate dependence of strength. If other mechanisms that lead to dynamic weakening operate, such as reduction in effective normal stress during seismic slip, then the stress drop indicated in Figure 10 is an underestimate.

## CONCLUSIONS

1) The recent advances in our understanding of friction constitutive relations and of frictional slip mechanisms allows the construction of more complete rheologic descriptions of the crust. These new rheologic descriptions should be able to address questions of crustal strength and the stability of crustal faults.

2) The weakening effect of water on faults at elevated temperatures and low rates demonstrated in experiments is due to the activation of solution-transfer processes. The behavior of faults within this water-weakened regime can be described with friction (pressure-dependent) constitutive relations.

3) Solution-transfer assisted frictional slip mechanisms are probably important within the regime covering the transition from brittle, frictional faulting to viscous, dislocation creep. Further

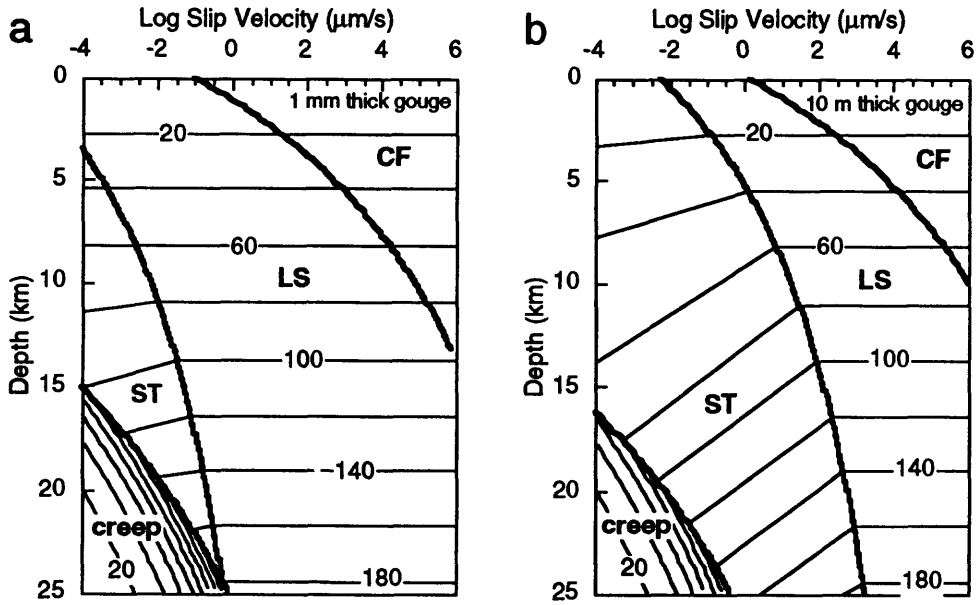



Figure 9. Dominant friction mechanism and shear strength (MPa) of the San Andreas fault as a function of depth and slip rate for (a) 1-mm thick and (b) 10-m thick gouge zones based on the rheologic model for wet quartz. Thick lines denote the boundaries between mechanism fields. The LS mechanism is rate-weakening and represents the region of potentially unstable slip.

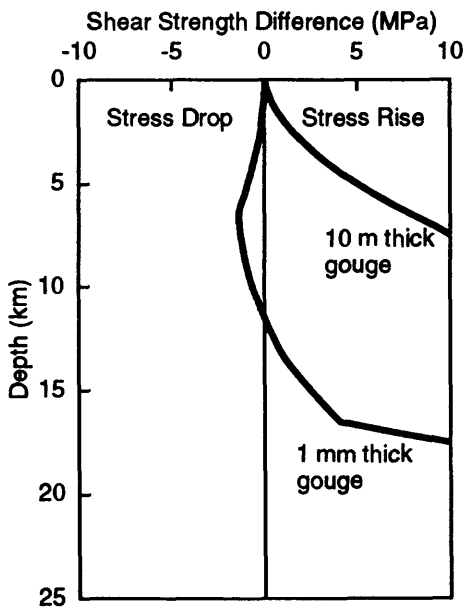



Figure 10. Difference in frictional strength of the San Andreas fault at the average slip rate of 30 mm/yr and at a coseismic slip rate of 1 m/s for the cases of a 1 mm and 10 m thick gouge zone based on the rheologic model of wet quartz. See text for further description.

study of mixed mechanism deformation processes involving combined cataclasis, plasticity and solution-transfer processes are necessary for a accurate characterization of crustal rheology.

4) The transition from friction mechanisms leading to the rate-weakening behavior necessary for slip instability to mechanisms producing rate-strengthening behavior depends on depth (temperature), slip rate, and gouge thickness. At geologic slip rates, the strength of faults is decreased and the stability is increased with an increase in gouge thickness. Only very thin faults display the rate-dependent characteristics necessary for nucleation of seismic slip throughout the upper crust (8 to 12 km depth).

ACKNOWLEDGEMENTS

J.S. Chester and R.H. Sibson provided comments on the manuscript, and H-F. Wang provided assistance in generating the deformation mechanism maps. This research was supported by the U.S. Geological Survey (USGS), Department of Interior, under awards 14-08-0001-G1768 and 14-92-G-2203.

REFERENCES

Biegel, R.L., C.G. Sammis, and J.H. Dieterich, The frictional properties of a simulated gouge having a fractal particle size distribution, *J. Struct. Geol.*, *11*, 827-846, 1989.

Blanpied, M.L., D.A. Lockner, and J.D. Byerlee, Fault stability inferred from granite sliding experiments at hydrothermal conditions, *Geophys. Res. Lett.*, *18*, 609-612, 1991.

Brace, W.F., and J.D. Byerlee, Stick-slip as a mechanism for earthquakes, *Science*, *153*, 990-992, 1966.

Brace, W.F., and D.L. Kohlstedt, Limits on lithospheric stress imposed by laboratory experiments, *J. Geophys. Res.*, *85*, 6248-6252, 1980.

Byerlee, J.D., Friction in rocks, *Pure Appl. Geophys.*, *116*, 615-626, 1978.

Chester, F.M., The brittle-ductile transition in a deformation-mechanism map for halite, *Tectonophysics*, *154*, 125-136, 1988.

Chester, F.M., Dynamic recrystallization in semi-brittle faults, *J. Struct. Geol.*, *11*, 847-858, 1989.

Chester, F.M., Structures in exhumed fault zones and implications for the dynamics of earthquake nucleation and slip, *Eos Trans. AGU*, *74*, 295, 1993.

Chester, F.M., Effects of temperature on friction: constitutive equations and experiments with quartz gouge, *J. Geophys. Res.*, in press, 1994.

Chester, F.M., J.P. Evans, and R.L. Biegel, Internal structure and weakening mechanisms of the San Andreas fault, *J. Geophys. Res.*, *98*, 771-786, 1993.

Chester, F.M. and N.G. Higgs, Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions, *J. Geophys. Res.*, *97*, 1859-1870, 1992.

Chester, F.M., and J.M. Logan, Frictional faulting in polycrystalline halite: correlation of microstructure, mechanisms of slip and constitutive behavior, in *The Brittle-Ductile Transition in Rocks: The Heard Volume*, *Geophys. Monogr. Ser.*, vol. 56, edited by A.G. Duba et al., pp. 49-65, AGU, Washington, D. C., 1990.

Davis, D., J. Suppe, and F.A. Dahlen, Mechanics of fold-and-thrust belts and accretionary wedges, *J. Geophys. Res.*, *88*, 1153-1172, 1983.

Dieterich, J.H., Time-dependent friction and the mechanics of stick-slip, *Pure Appl. Geophys.*, *116*, 790-805, 1978.

Dieterich, J.H., Modeling of rock friction, 1, experimental results and constitutive equations, *J. Geophys. Res.*, *84*, 2161-2168, 1979.

Dieterich, J.H., Constitutive properties of faults with simulated gouge, in *Mechanical Behavior of Crustal Rocks: The Handin Volume*, *Geophys. Monogr. Ser.*, vol. 24, edited by N.L. Carter et al., pp. 103-120, AGU, Washington, D. C., 1981.

Dula, W.F., Jr., High temperature deformation of wet and dry artificial quartz gouge, Ph.D. Dissertation, 382 pp., Texas A & M University, College Station, TX, 1985.

Frost, H.J., and M.F. Ashby, *Deformation-mechanism maps: the plasticity and creep of metals and ceramics*, Pergamon Press, Oxford, 166 pp., 1982.

Goetze, C., and B. Evans, Stress and temperature in the bending lithosphere as constrained by experimental rock mechanics, *Geophys. J. R. astr. Soc.*, *59*, 463-478, 1979.
Gu, J.-C., J.R. Rice, A.L. Ruina, and S.T. Tse, Slip motion and stability of a single degree of freedom elastic system with rate and state dependence friction, *J. Mech. Phys. Solids*, *32*, 167-196, 1984.
Higgs, N.G., Mechanical properties of ultrafine quartz, chlorite and bentonite in environments appropriate to upper-crustal earthquakes, Ph.D. Dissertation, 267 pp., Texas A & M University, College Station, 1981.
Kirby, S.H., Rheology of the lithosphere, *Rev. Geophys. Space Phys.*, *21*, 1458-1487, 1983.
Koch, P. S., J. M. Christie, A. Ord, and R. P. George, Jr., Effect of water on the rheology of experimentally deformed quartzite, *J. Geophys. Res.*, *94*, 13975-13996, 1989.
Lachenbruch, A. H., and A. McGarr, Stress and heat flow, in The San Andreas fault system, California, edited by R. E. Wallace, pp. 261-277, *U.S. Geol. Surv. Prof. Pap.*, *1515*, 1990.
Lockner, D.A., Summers, R., and J.D. Byerlee, Effects of temperature and sliding rate on frictional strength of granite, *Pure Appl. Geophys.*, *124*, 445-469, 1986.
Marone, C., and B. Kilgore, Scaling of the critical slip distance for seismic faulting with shear strain in fault zones, *Nature*, *362*, 618-621, 1993.
Marone, C., C.B. Raleigh, and C.H. Scholz, Frictional behavior and constitutive modeling of simulated fault gouge, *J. Geophys. Res.*, *95*, 7007-7025, 1990.
Marone, C., and C.H. Scholz, Particle-size distribution and microstructures within simulated fault gouge, *J. Struct. Geol.*, *11*, 799-814, 1989.
Poirier, J-P., *Creep of Crystals: High Temperature Deformation Processes in Metals, Ceramics and Minerals*, 260 pp., Cambridge University Press, New York, 1985.
Ruina, A.L., Slip instability and state variable friction laws, *J. Geophys. Res.*, *88*, 10,359-10,370, 1983.
Rutter, E.H., On the nomenclature of mode of failure transitions in rocks, *Tectonophysics* , *122*, 381-387, 1986.
Rutter, E.H., and D.H. Mainprice, The effect of water on the stress relaxation of faulted and unfaulted sandstone, *Pure Appl. Geophys.*, *116*, 634-654, 1978.
Rutter, E.H., and D.H. Mainprice, On the possibility of slow fault slip controlled by a diffusive mass transfer process, *Gerlands Beitr. Geophys.*, *88*, 154-162, 1979.
Rutter, E.H., and S.H. White, The microstructures and rheology of fault gouges produced experimentally under wet and dry conditions, 300 to 400°C, *Bull. Mineral.*, *102*, 101-109, 1979.
Sammis, C.G., Relating fault stability to fault-zone structure, *Eos Trans. AGU*, *74*, 295-296, 1993.
Sammis, C., G. King, and R. Biegel, The kinematics of gouge deformation, *Pure Appl. Geophys.*, *125*, 777-812, 1987.
Scholz, C.H., *The Mechanics of Earthquakes and Faulting*, 439 pp., Cambridge University Press, New York, 1990.
Sibson, R.H., Frictional constraints on thrust, wrench and normal faults, *Nature*, *249*, 542-544, 1974.
Sibson, R.H., Continental fault structure and the shallow earthquake source, *J. Geol. Soc. London*, *140*, 741-767, 1983.
Sibson, R.H., Roughness at the base of the seismogenic zone: Contributing factors, *J. Geophys. Res.*, *89*, 5791-5799, 1984.
Tse, S.T., and J.R. Rice, Crustal earthquake instability in relation to the depth variation of frictional slip properties, *J. Geophys. Res.*, *91*, 9452-9472, 1986.

# LOW FRICTION DURING SLIDING ON SIMULATED FAULTS IN POROUS QUARTZ SANDSTONE AT HYDROTHERMAL CONDITIONS

Stephen F Cox
Research School of Earth Sciences, The Australian National University, Canberra, ACT 0200, Australia

## ABSTRACT

A suite of high temperature experiments has been used to investigate frictional sliding behavior in the presence of chemically active pore fluids. Experiments have been conducted on bare interfaces in porous quartz sandstone at 1200K, a confining pressure of 300 MPa, and at a pore fluid pressure of 200 MPa. Sliding interfaces were inclined at $30^{o}$ to the axial shortening direction, and shear displacement rates ranged between 0.2 $\mu m\ s^{-1}$ and 1 $\mu m\ s^{-1}$. Argon and water have been used as pore fluids. At the test conditions the argon pore fluid is unreactive, whereas the presence of pore water promotes extremely rapid dissolution-precipitation reactions.

In the presence of argon, and at an imposed displacement rate of 0.7 $\mu m\ s^{-1}$, simulated faults exhibit slip hardening behavior. Slip commenced at shear stresses around 60 MPa and at friction coefficients around 0.5.

The presence of pore water has a spectacular effect on the mechanical behavior of sliding interfaces. At a displacement rate of 0.8 $\mu m\ s^{-1}$, the initiation of sliding at shear stresses around 50 MPa was accompanied by a rapid slip event and decrease in shear resistance. During ensuing stable sliding, the coefficient of friction had a value around 0.3. A transition to stick-slip behavior has been found at lower displacement rates.

Microstructural observations indicate that dry sliding over distances up to 1mm produced only minor surface damage. However, during wet sliding, the slip surfaces have been extensively modified by dissolution-precipitation processes. Wear damage that has developed early during sliding has been etched to form well-developed slickenlines over the entire slip surface, even after less than 1mm of displacement.

The low friction coefficients measured during high temperature wet sliding are substantially lower than values found previously for simulated bare faults in quartz-rich rocks at lower temperatures. This result, together with the microstructural observations, suggests the possibility that the presence of reactive pore fluids can play a significant role in controlling fault stability and decreasing resistance to frictional sliding in faults at seismogenic depths in the Earth's continental crust.

## INTRODUCTION

Numerous studies of exhumed, ancient fault zones indicate that active faults can be sites of enhanced fluid activity, and that fluid-rock reactions can play a key role in modifying the internal structure of faults. Major effects include repeated episodes of interseismic hydrothermal sealing and alteration of faults and fault products (eg Stel, 1981; House and Gray, 1982; Sibson et al, 1988; Cox et al, 1991; Evans, 1990; Chester et al, 1993; Cox, this volume). The effects of progressive chemical fluid-rock interaction between seismic slip events or during fault creep are likely to cause substantial, time-dependent changes in the

mechanical behavior and stability of fault zones. There is a clear need for our developing understanding of fault mechanics and the factors controlling earthquake nucleation and recurrence to be underpinned by experimental investigations of the role of chemically active pore fluids in controlling fault behavior.

This paper presents the results of a suite of experiments which has been used to examine sliding behavior on simulated faults in quartz sandstone at elevated temperatures and in the presence of reactive pore fluids. The experiments demonstrate a substantial influence of dissolution-precipitation reactions on the strength, mechanical behavior and structure of sliding interfaces. In particular, the study provides new insights about the role of chemically active pore fluids in controlling fault strength and stability at seismogenic depths.

## EXPERIMENTAL TECHNIQUES AND CONDITIONS

Sliding experiments have been performed at confining pressures of 300 MPa, pore pressures of 200 MPa, and at temperatures of 1200K in an argon confining medium, triaxial testing apparatus described by Cox & Paterson (1991). Tests have been conducted at nominally constant displacement rates between 0.2 $\mu ms^{-1}$ and 1 $\mu ms^{-1}$. Experimental constraints have limited maximum shear displacements to 1.3 mm.

The test specimens consist of two cylindrical forcing blocks that have been fabricated from a quartz sandstone (Fontainbleu Sandstone). The sliding interface on the forcing blocks is inclined at $30^{o}$ to the maximum compression direction, and has been prepared by grinding on a 400 grit diamond lap. The test piece has a diameter of 9.97 mm and a length between 24 mm and 27.5 mm. The sandstone has a porosity of 10-12%, and is composed of equant quartz grains with an average grainsize of 0.25 mm. Tests on comparable samples of Fontainbleu Sandstone (Bourbie and Zinszner, 1985) indicate that the test pieces have a permeability of around $10^{-13}$ $m^2$.

The samples were contained within a thin-walled copper sleeve and an outer iron jacket, using an experimental arrangement described by Cox and Paterson (1991). At the completion of deformation tests, samples were extracted from their metal jackets by dissolution of the iron and copper jacketting materials in concentrated nitric acid. Microstructural characterisation of interfaces has been performed by SEM secondary electron imaging of gold-coated sample surfaces.

## EXPERIMENTAL RESULTS

*Mechanical Data*

At an imposed shear displacement rate of 0.7 $\mu m\ s^{-1}$ at 1200K, and in the presence of argon pore fluid, slip commences at shear stresses around 60 MPa (Fig 1a). Further displacement is characterised by slip hardening. The coefficient of friction has a value around 0.5. This is significantly lower than values between 0.6 and 0.8 determined in the same experimental arrangement, and for the same effective pressure at room temperature.

The introduction of pore water at elevated temperatures has a spectacular effect on the mechanical behaviour of the sliding interfaces (Fig. 1b). At an imposed sliding rate of 0.8 $\mu ms^{-1}$, initial build-up of shear stress is terminated by rapid slip at shear stresses around 50 MPa. Shear stress relief is followed by stable sliding in which shear strength has a weak positive dependence on sliding rate. During stable sliding the coefficient of friction has a value around 0.3. A transition to stick-slip behavior has been found at lower displacement rates (Fig 1c).

*Microstructures*

The machined sliding surfaces in the starting material are characterised by the presence of (1) fine wear grooves with widths up to about 0.5 $\mu m$ and depths less than 0.2 $\mu m$, and (2) much more irregularly-shaped pits that are up to 10 $\mu m$ in diameter and several

microns deep. The latter structures appear to have developed by surface plucking during grinding. Also present are much larger, primary intergranular pores which have diameters up to 200 μm.

During high temperature sliding experiments, the wear grooves that were formed during specimen preparation have been oriented at about 30° to the slip direction. Following high temperature slip in the presence of argon pore fluid, some of the initial wear grooves, and associated irregular pits, tend to be slightly more deeply etched than in the starting material. The sliding surfaces also contain a new set of very fine wear grooves parallel to the experimental slip direction (Fig. 2). Very little surface debris is present, although some pores contain some fine (typically < 1 μm) angular debris that has presumably been derived from surface wear.

The presence of pore water has a dramatic effect on the surface structure that evolves during isostatic hydrothermal treatment prior to sliding, as well as during sliding itself. Specimens typically have been held at the test pressure and temperature for one hour prior to the commencement of sliding. During this isostatic hydrothermal treatment, the initial (specimen preparation) wear grooves, as well as the rest of the sliding surface, become substantially modified by dissolution-precipitation reactions (Fig. 3). Individual wear grooves can be etched out to widths around 10 μm and up to several microns deep. Away from recognisable grooves, dissolution-precipitation processes have generated an irregular topography with an amplitude and wavelength of several microns. Euhedral quartz crystal faces are locally developed, as are ridge and channel structures similar to those described by Cox and Paterson (1991).

After the initial rapid slip increment, the slip surface is covered by a set of new grooves that are parallel to the slip direction. The original (pre-test) wear grooves have been largely obliterated. During continued sliding the newly-formed grooves become progressively deeper due to continued etching by the pore fluid. Dissolution-precipitation reactions also further modify the fine-scale structure within and around the new grooves (Fig. 4). The experimentally-produced slickenlines have maximum lengths around 150 μm. This is comparable to the magnitude of displacement associated with initial rapid slip prior to stable sliding (Fig. 1b), and suggests that slickenlines may not have increased in length during continued 'wet' sliding. Interestingly, there appears to be no significant difference in the structure of etched wear tracks developed during stable sliding and stick-slip sliding.

## DISCUSSION AND IMPLICATIONS

The experiments provide a clear demonstration of the importance of reactive pore fluids in modifying the mechanical behavior and resistance to sliding on bare rock interfaces. The presence of pore water at elevated temperatures has been found to be associated with a substantial reduction in the resistance to sliding when compared with the frictional behavior at the same conditions, but in the presence of an inert pore fluid.

Blanpied et al (1992) have previously demonstrated reduced sliding resistance on simulated, gouge-bearing faults at elevated temperatures and high water pressures. In their case, decreased sliding resistance was ascribed to the development of abnormally high and uncontrolled fluid pressures in the simulated fault. This situation arose as a result of hydrothermal sealing isolating the gouge layer from the pore fluid reservoir during their experiment. Progressive compaction of the gouge then caused a gradual increase in fluid pressure and decrease in effective normal stress in the simulated fault.

For the experiments described in this paper, the high permeability of the slider blocks and the asperity structure on the clean sliding interfaces apparently has ensured good fluid connectivity between the simulated fault and the external pore fluid reservoir. Accordingly, the observed decrease in sliding resistance is ascribed to an intrinsically low friction coefficient, rather than the development of high, uncontrolled fluid pressures and a consequent decreased effective normal stress across the sliding interface.

The friction coefficients measured in these high temperature hydrothermal experiments are substantially lower than values found previously in quartz-rich rocks at substantially lower temperatures where water is not particularly reactive. The experimental demonstration that friction coefficients can be as low as 0.3 in quartz sandstone raises the possibility that other silicate minerals may also have intrinsically low friction coefficients in the presence of chemically active pore fluids. If so, active crustal faults that are fluid channelways may be substantially weaker than predicted by frictional sliding experiments conducted in the presence of non-reactive pore fluids.

The microstructure of the sliding surfaces clearly illustrates the importance of dissolution-precipitation processes in rapidly modifying the asperity structure of the simulated faults. The substantially reduced frictional resistance in this environment is interpreted to have been influenced by thermally-activated dissolution-precipitation processes producing a continuously evolving asperity structure during sliding. However, thermally-activated stress corrosion cracking processes and direct lubrication of sliding interfaces by an adsorbed fluid layer may also have had a role in reducing resistance to sliding.

The results presented in this study highlight one way in which reactive pore fluids may have a substantial influence on fault strength and stability. Widespread evidence for the presence of hydrothermal fluids in active faults at seismogenic depths requires that a realistic basis for understanding earthquake nucleation and recurrence will need to be underpinned by both laboratory-based and field-based inverstigations of processes whereby chemical fluid-rock interactions can influence the mechanical behavior of fluid-bearing fault zones on geologic time-scales.

## Acknowledgements

The Fontainbleu Sandstone samples were kindly provided by J.-P. Poirier and M. Zamorra of the Institut de Physique du Globe, Paris. Mervyn Paterson, Rick Sibson and Simon Cox have provided much stimulating discussion. Experimental facilities have been ably maintained by Graeme Horwood. The assistance of the staff at the ANU Electron Microscopy Unit is gratefully acknowledged. This study was conducted with the support of an ARC Senior Research Fellowship. The Director, RSES, is acknowledged for continuing support of this research. The USGS generously provided financial support allowing the author to participate in the Fish Camp workshop.

## References

Blanpied, M. L., Lockner, D. A. and Byerlee, J. D., 1992. An earthquake mechanism based on rapid sealing of faults. *Nature* **358**, 574-576.

Bourbie, T. and Zinszner, B., 1985. Hydraulic and acoustic properties as a function of porosity in Fontainbleu Sandstone. *J .Geophys. Res.* **90**, 11524-11532.

Chester, F.M., Evans, J.P. and Biegel, R.L., 1993. Internal structure and weakening mechanisms of the San Andreas Fault. *J. Geophys. Res.* **98**, 771 - 786.

Cox, S. F. and Paterson, M. S., 1991. Experimental dissolution-precipitation creep in quartz aggregates at high temperatures. *Geophys. Res. Lett.* **18**, 1401-1404.

Cox, S. F., Wall, V. J., Etheridge, M. A., and Potter, T. F., 1991. Deformational and metamorphic processes in the formation of mesothermal vein-hosted gold deposits examples from the Lachlan Fold Belt in central Victoria. *Ore Geology Reviews* **6**, 391- 423.

Evans, J.P., 1990. Textures, deformation mechanisms, and the role of fluids in the cataclastic deformation of granitic rocks. In: Knipe, R.J. and Rutter, E.H. (eds), "Deformation Mechanisms, Rheology and Tectonics". *Geol. Soc Special Publn.* **54**, pp 29 -39.

House, W.M. and Gray, D.R., 1982. Cataclasites along the Saltville Thrust, U.S.A. and their implications for thrust sheet emplacement. *J. Struct. Geol.* **4**, 257 - 269.
Sibson, R. H., Robert, F. and Poulson, K. H., 1988. High-angle reverse faults, fluid pressure cycling and mesothermal gold deposits. *Geology* **16**, 551- 555.
Stel, H., 1981. Crystal growth in cataclasites: Diagnostic microstructures and implications. *Tectonophysics* 78, 585 - 600.

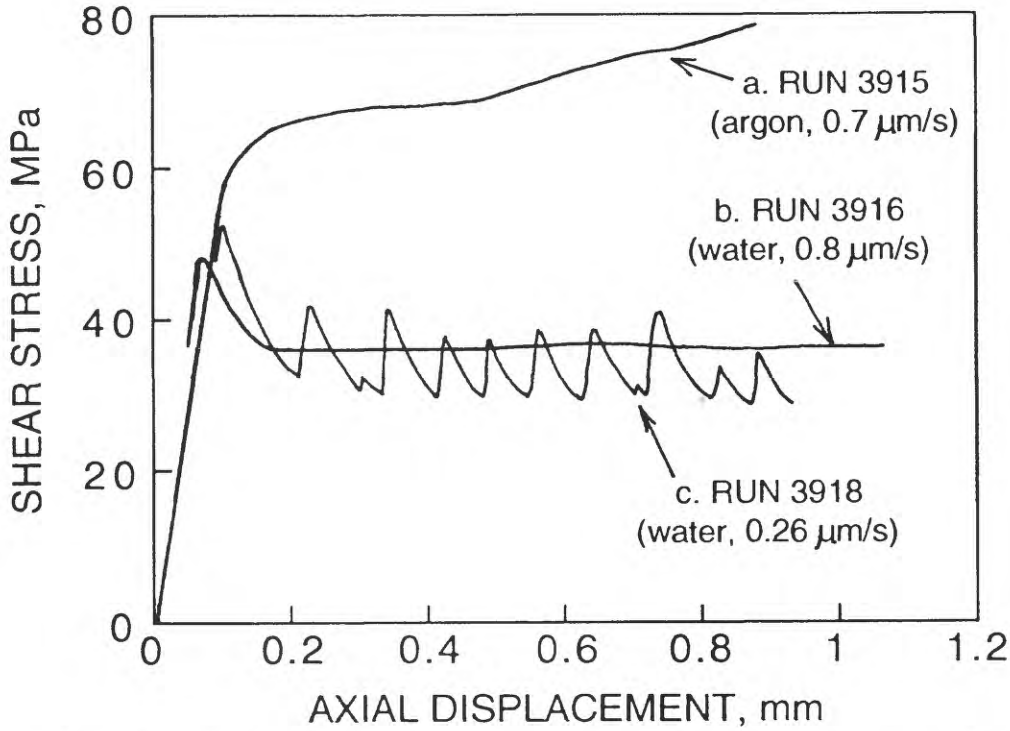



**Figure 1.** Shear stress - displacement curves for sliding experiments on Fontainbleu Sandstone at 1200K, 300MPa confining pressure, and 200 MPa pore pressure. **a**. Argon pore fluid, shear displacement rate of 0.7 μm $s^{-1}$. **b**. Water pore fluid, shear displacement rate of 0.8 μm $s^{-1}$. **c**. Water pore fluid, shear displacement rate of 0.26 μm $s^{-1}$.

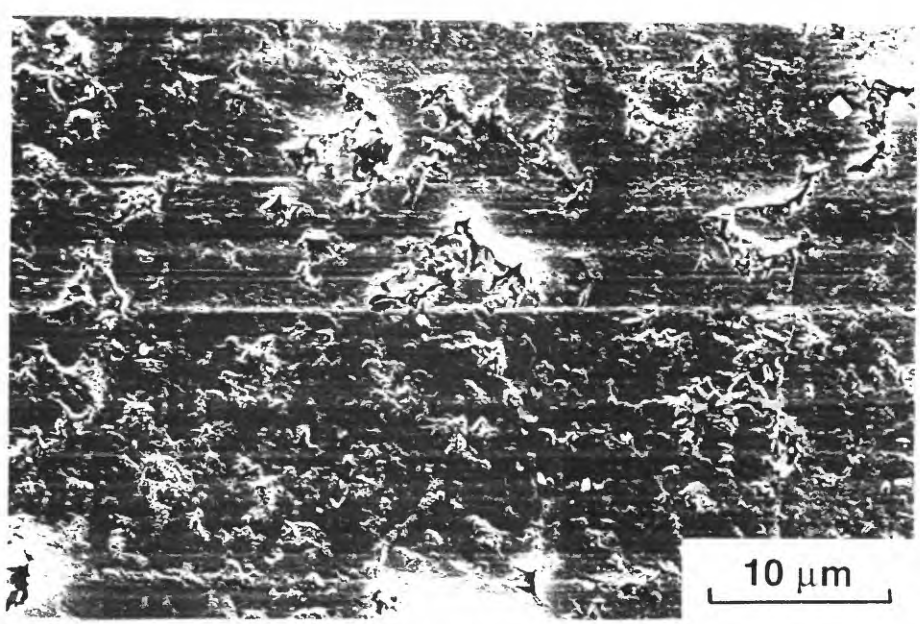



**Figure 2.** SEM micrograph illustrating the microstructure on the surface of a Fontainbleu Sandstone slider block that has slipped 1.0 mm at 1200K, 300 MPa confining pressure, and 200 MPa argon pore fluid pressure. Fine wear tracks have formed parallel to the slip direction. Many of the small, irregularly-shaped pits have apparently formed during specimen preparation.

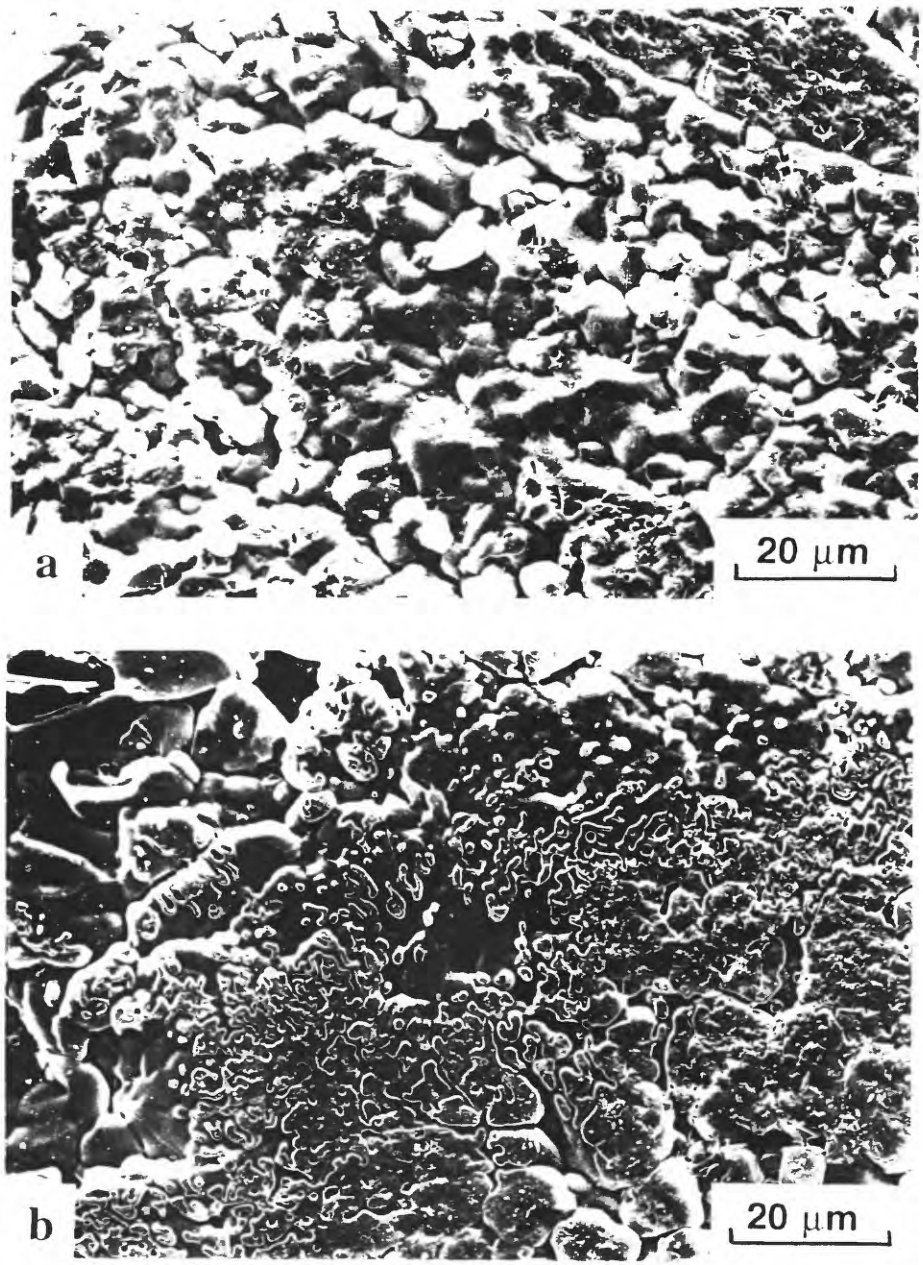



**Figure 3.** Microstructure of the surface of a slider block after one hour of isostatic hydrothermal treatment at 1200K, 300 MPa confining pressure, and 200 MPa pore water pressure.
**a.** Typical area of irregular topography. **b.** Area with well-developed, lobate ridge and channel structures.

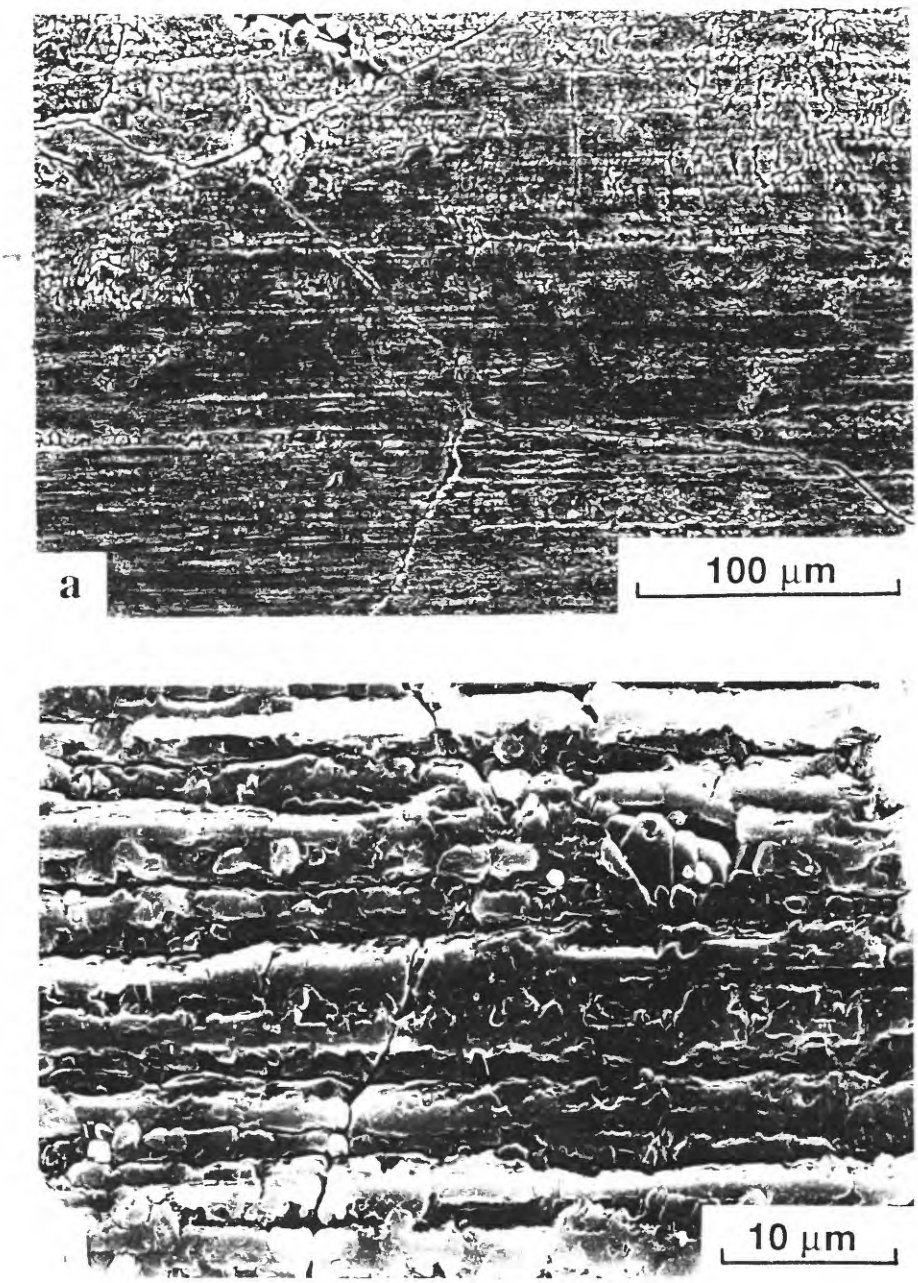



**Figure 4.** SEM micrographs illustrating the microstructure on a Fontainbleu Sandstone interface after just over 1mm of slip at 1200K, 300 MPa confining pressure, and 200 MPa pore water pressure. Surface damage, which has been selectively etched by the pore fluid during slip, has produced closely spaced slickenlines that are continuous for lengths up to 150 μm. Despite the total slip being so small, the new microstructure has totally obliterated the pre-existing surface structure that was developed during specimen preparation and hydrothermal treatment prior to sliding. The slip direction is parallel to the long axis of the micrographs.

**a.** Typical structure of slickenlines on the surface of a slider block.

**b.** Detail of the microstructure of the etched wear tracks. Note the cluster of euhedral quartz crystals near the top right part of the micrograph.

# THE EFFECT OF PORE FLUID CHEMISTY ON THE FRICTION OF QUARTZ GOUGE

*Terry E. Tullis*
*Department of Geological Sciences, Brown University, Providence, RI 02912*

Surfaces in stationary contact increase in strength with time; this is commonly referred to as the evolution effect. It is the most important aspect of rock friction to understand, because it is responsible for the restrengthening of surfaces between slip episodes. Without such restrengthening, repeated unstable slip is impossible. The most straightforward demonstration of the evolution effect is the slide-hold-slide test. In such a test, sliding takes place at constant velocity until the sliding resistance attains a steady-state value. Sliding is then stopped for some time, after which sliding is resumed. The resistance climbs to a peak value larger than the steady state-value prior to the hold, subsequently decaying to the original steady-state value. There is a nearly linear relationship between the magnitude of the peak value and the log of the hold time, over at least 4-5 orders of magnitude (*Dieterich*, 1972).

The evolution effect also is responsible for a tendency for frictional resistance to be lower at higher sliding velocities. The lifetime of a frictional contact will be inversely proportional to the sliding velocity: the lower the velocity is, the longer will be the duration of a given contact. Just as in the slide-hold-slide tests, the older a contact is, the stronger it will be before it breaks. Thus, steady sliding can have higher strength at lower velocity, the reverse of the usual observation of velocity dependence of deformation strength. After abrupt changes in velocity a new steady-state resistance is established over a characteristic slip distance. This has been interpreted (*Dieterich*, 1978, 1979) as the distance required to destroy all the contacts established at one velocity and to create a new set of contacts having an average age appropriate to the new velocity.

We still do not know the reason for the strengthening with time and do not know how to extrapolate them to longer times than are attainable in the lab. It must be stressed that this extrapolation is a severe problem, and just as in the case of ductile rock deformation we need to know the processes. Laboratory hold tests have been conducted for times from $10^0$ to $10^5$ seconds, but for great earthquakes along the San Andreas fault the inter-event times are about $10^{9.7}$ seconds.

The experiments of *Dieterich and Conrad* (1984) that show that the evolution effect is eliminated if the sliding surfaces are dried thoroughly have led us to a hypothesis on the origin of the evolution effect. In the experiments of Dieterich and Conrad, re-introduction of water vapor causes the return of the evolution effect. Structured water is attracted to broken surfaces in quartz and other silicates (*Parks*, 1984), in part related to their negative charge. Because this adsorbed water is quite tightly bonded to most silicate surfaces, it will prevent the close approach of the surfaces, thereby inhibiting bonding between the surfaces. We believe that the evolution effect may be caused by the gradual removal of contaminant, most likely water, from between contacting surfaces resulting in a stronger bond.

If this is true, then altering the chemical environment of a sliding surface should alter the frictional behavior in predictable ways. A solution containing electrolytes should make bonding between the surfaces even more difficult than in a neutral pH pure water solution, because the cations become hydrolyzed and attracted to the negatively charged surface, inhibiting the close approach of the surfaces (*Israelachvili*, 1992). The electrolytes represent a greater contamination of the surface than pure water and should lead to a greater evolution effect. We postulated that experiments conducted at a pH of about 2.6, the so-called "point of zero charge" for quartz (*Jed-*

*nacak et al.*, 1974; *Chibowski*, 1979; *Parks, 1984*), would exhibit little or no evolution effect. At that pH there is a sufficient concentration of protons that they imbed themselves in the surface, cancelling the surface charge and reducing the tendency for water or hydrolyzed cations to be attracted to the surface (*Pashley*, 1985). This would in turn render the contaminant layer ineffective, allowing direct bonding of the silicate surfaces.

We have recently completed some preliminary experiments that support these predictions, although their unambiguous interpretation must await further experiments. The results are shown in Figure 1 for three ten-fold velocity decreases at an identical displacement in three separate experiments using 1 mm layers of ultra-fine quartz powder. In Figure 1a we show a control experiment which is done in atmospheric humidity like the control experiment of *Dieterich and Conrad* (1984). This experiment contained no liquid water. It shows the typical evolution effect as the decay back up toward the final steady-state level, following the drop due to the direct effect response to the velocity decrease. An experiment done with a solution of nitric acid at an initial pH of 2.6 is shown in Figure 1b. As predicted, it shows no evolution effect. Finally in Figure 1c we show the results from an experiment buffered at a pH of 11 using a buffer that includes NaOH and NaCl. This experiment shows a larger evolution effect than the control experiment. The enhanced magnitude of the evolution effect in the pH 11 experiment is better shown in Figure 2a which shows the results of slide-hold-slide tests. It is clear from the figure that the strengthening with time is larger for this experiment than any of the others.

As might be expected with only a few experiments so far, there are complications. The results of the pH 2.6 experiment are not completely unambiguous. Only in the first two velocity changes did the evolution effect appear to be absent (Figure 1b). Subsequent velocity changes and slide-hold-slide tests (Figure 2a) showed a typical magnitude for the evolution effect. We believe that the supply of $H^+$ ions in the solution may have been exhausted early in the experiment, shifting the pH toward more neutral values. The pH in this experiment was not buffered at a pH of 2.6; the concentration of $HNO_3$ was simply chosen to give that value. Nitric acid completely dissociates, and so as $H^+$ ions are removed from solution by the negatively charged surface of the quartz, it depletes the supply causing the pH to increase away from the point of zero charge. An approximate calculation of the number of $H^+$ ions available in the solution, given the porosity of gouge and the surface area of the fine-grained quartz particles in the gouge, suggests that the number of $H^+$ ions was about the right order of magnitude to be completely used up by adsorbing onto the surfaces. Thus there would be no reserve supply of $H^+$ ions for the greatly increased surface area generated by comminution early in the experiment.

The experiment at pH 11 is also not completely unambiguous. Two possible explanations could be offered for the greater evolution effect in this experiment. The buffered solution contains both NaCl and NaOH, and the presence of $Na^+$ ions in the solution may enhance the evolution effect independently of the high pH, as explained above. However, the pH 11 experiment was actually planned originally because the solubility of $SiO_2$ is 16 times higher at a pH of 11. If pressure solution is involved in the evolution effect, then high pH might enhance the effect by providing a flux of dissolved material 16 times greater.

We believe that the enhanced evolution effect is caused by the presence of the $Na^+$ ions in solution. If the greater solubility of silica were responsible, we might expect an increase by a factor of 16 rather than the observed factor of about 2. Furthermore, we see three other effects only at low velocities or long hold times that we believe are caused by the increased solubility. First, at the longest hold time in Figure 2a the strength upon reloading was considerably below the linear trend.

Second, after the longest relaxation times, the steady state level after the reload is much lower than the level before the relaxation is started, as shown in Figure 2b. For all of the other experiments this steady state level is reattained after the holds. Something seems to have occurred in the two longest holds in this experiment to cause the steady state strength to decline. Third, the form of the relaxation indicates that a new process becomes important at long times, and therefore at very slow velocities. Figure 2c shows three curves of normalized frictional strength versus the log of relaxation time. The normalized stress declines in a relatively straight line for all of the experiments except the alkaline one which shows a dramatic reduction in slope after about 10,000 seconds (2.8 hours). This form of the relaxation curve is similar to that seen by *Chester and Higgs* (1992) in quartz experiments done under hydrothermal conditions and by *Reinen* (1993) in room temperature experiments on serpentine. In both of these cases the authors determined that a flow process was occurring at the lowest velocities, pressure solution for the quartz under hydrothermal conditions and probably dislocation glide for the serpentine. In our experiment at pH 11, we presume that pressure solution was responsible, due to the increased solubility of quartz.

While there are significant ambiguities in interpreting these experiments, the results are very exciting. We have demonstrated that we can make significant changes in the frictional behavior of quartz by changing the chemical environment of the sliding surface. We have shown several different effects including the possible elimination of the evolution effect by a low pH environment, the enhancement of the evolution (probably by introduction of electrolyte solution), and the appearance of a new process at low velocity and elevated pH. The results also suggest future experiments to resolve the ambiguities.

We have interpreted the observed elimination of the evolution effect at the point of zero charge in terms of the attachment of water to the surface as a contaminant. However, it is possible that the point of zero charge could also have an influence on the mobility of near surface dislocations by affecting their electronic core structure or charge (*Westwood et al.*, 1967). Additional experiments and analysis are needed to determine whether an unambiguous interpretation of the experimental data in terms of the process responsible for the evolution effect can be obtained. The roles played by various ions in affecting the near-surface dislocation mobility as described by *Westwood et al.* (1967) for MgO seems to involve the dipole moments of complex ions and organic molecules, and the addition of $CaCl_2$ to water did not alter the dislocation mobility. Dislocation mobility was increased by the adsorption of some species and decreased by the adsorption of others. Thus we anticipate that by using a variety of ions it should be possible to discriminate between their action as contaminants and as promoters of enhanced near-surface dislocation mobility.

## REFERENCES CITED


Chester, F.M., and Higgs, N.G., Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions, *J. Geophys. Res.*, *97*, 1859-1870, 1992.

Chibowski, E., Zeta potential and surface energy changes: quartz/n-heptane-water system, *J. Colloid Interface Sci*, *69*, 326-329, 1978.

Dieterich, J. H., Time-dependent friction in rocks, *J. Geophys. Res.*, *77*, 3690-3697, 1972.

Dieterich, J. H., Time-dependent friction and the mechanics of stick slip, *Pure Appl. Geophys.*, *116*, 790-806, 1978.

Dieterich, J. H., Modelling of rock friction, 1, Experimental results and constitutive equations, *J. Geophys. Res.*, *84*, 2161-2168, 1979.

Dieterich, J. H. and Conrad, G., Control of time- and velocity-dependent friction in rocks by adsorbed water, *J. Geophys. Res.*, *89*, 4196-4202, 1984.

Israelachvili, J. N., Adhesion forces between surfaces in liquids and condensable vapors, *Surface Science Reports, 14*, 109-159, 1992.

Jednacak, J, Pravdic, V., and Haller, W., The electrokinetic potential of glasses in aqueous electrolyte solutions, *J. Colloid Interface Sci., 49*, 16-23, 1974.

Parks, G. A., Surface and interfacial free energies of quartz, *J. Geophys. Res., 89*, 3997-4008, 1984.

Pashley, R.M., The effects of hydrated cation adsoprtion on surface forces between mica crystals and its relevance to colloidal systems, *Chemica Scripta, 25*, 22-27, 1985.

Reinen, L. A., *Frictional Behavior of Serpentenite: Experiments, Constitutive Model, and Implications for Natural Faults*. Ph.D. Thesis, Brown University, 177 p., 1993.

Westwood, A. R. C., Goldheim, D. L., and Lye, R. G., Rebinder effects in MgO, *Philos. Mag., 16*, 505-519, 1967.

## FIGURES

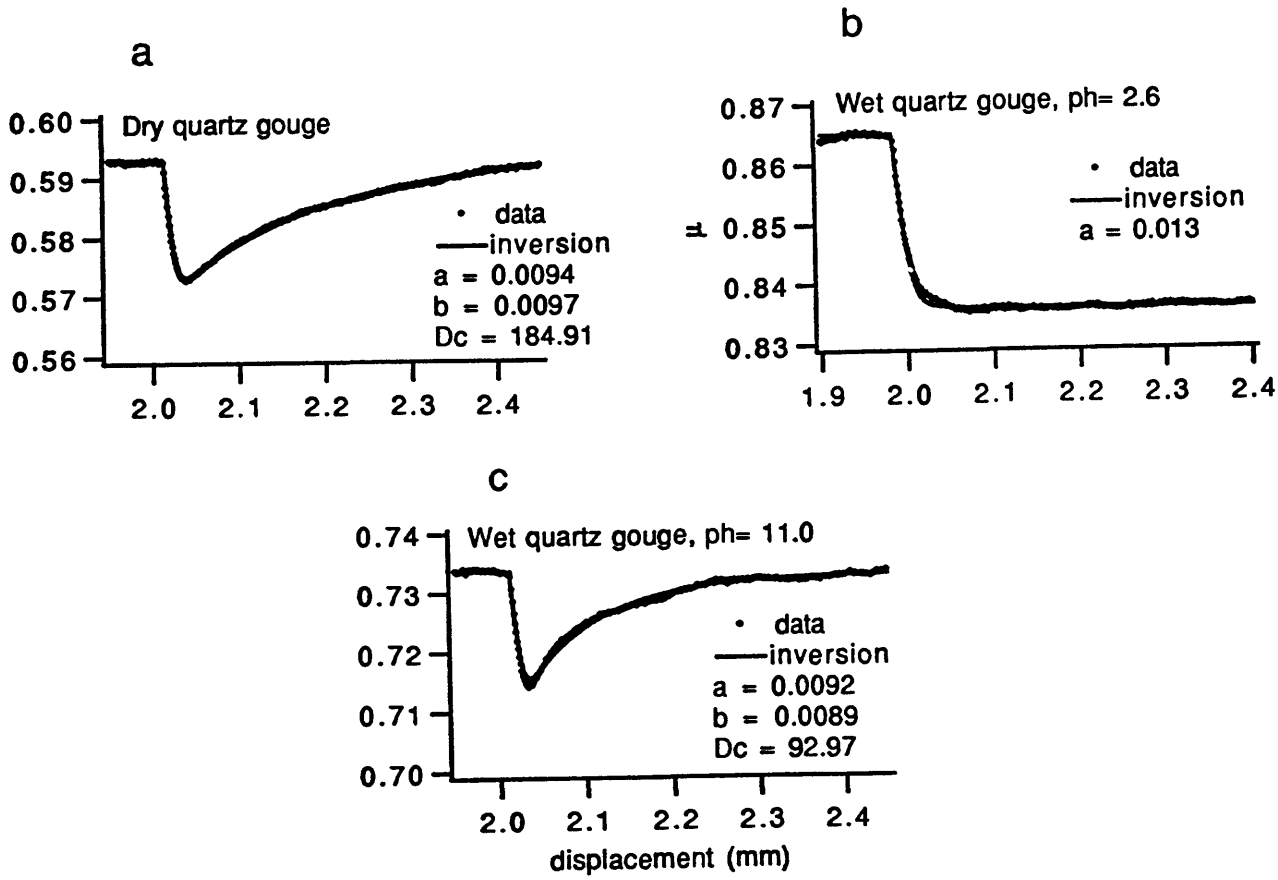



Figure 1. Comparison of frictional response to an decrease in loading velocity from 10.0 to 1.0 μm/sec for three layers of fine quartz gouge 1 mm thick, each run under different chemical environment. Although the fit is so good as to be virtually invisible, each trace is composed of a data set from an experiment (dots) and a numerical simulation using parameters determined by iterative least-squares inversion. The data are from very early in the runs (note the displacement scale) and include a substantial background trend that has been removed as part of the inversion. (A) Control experiment - room dry. This is a standard type response to a velocity step, showing a direct effect and evolution. In this case, the steady-state dependence on velocity is nearly neutral. Note the rather large value of $D_c$. (B) Nitric acid solution having a pH of 2.6. This is the point of zero charge, where the tendency of water to become bound to silicate surfaces will be eliminated or reduced due to saturation of the surface with protons. Thus the contaminant layer, which we believe may be responsible for the evolution effect, should not exist. In fact, the response is well fit with a model having no evolution effect. (C) Buffered solution having pH = 11.0. The magnitude of the evolution effect is not increased relative to the dry run, but $D_c$ is a factor of two smaller.

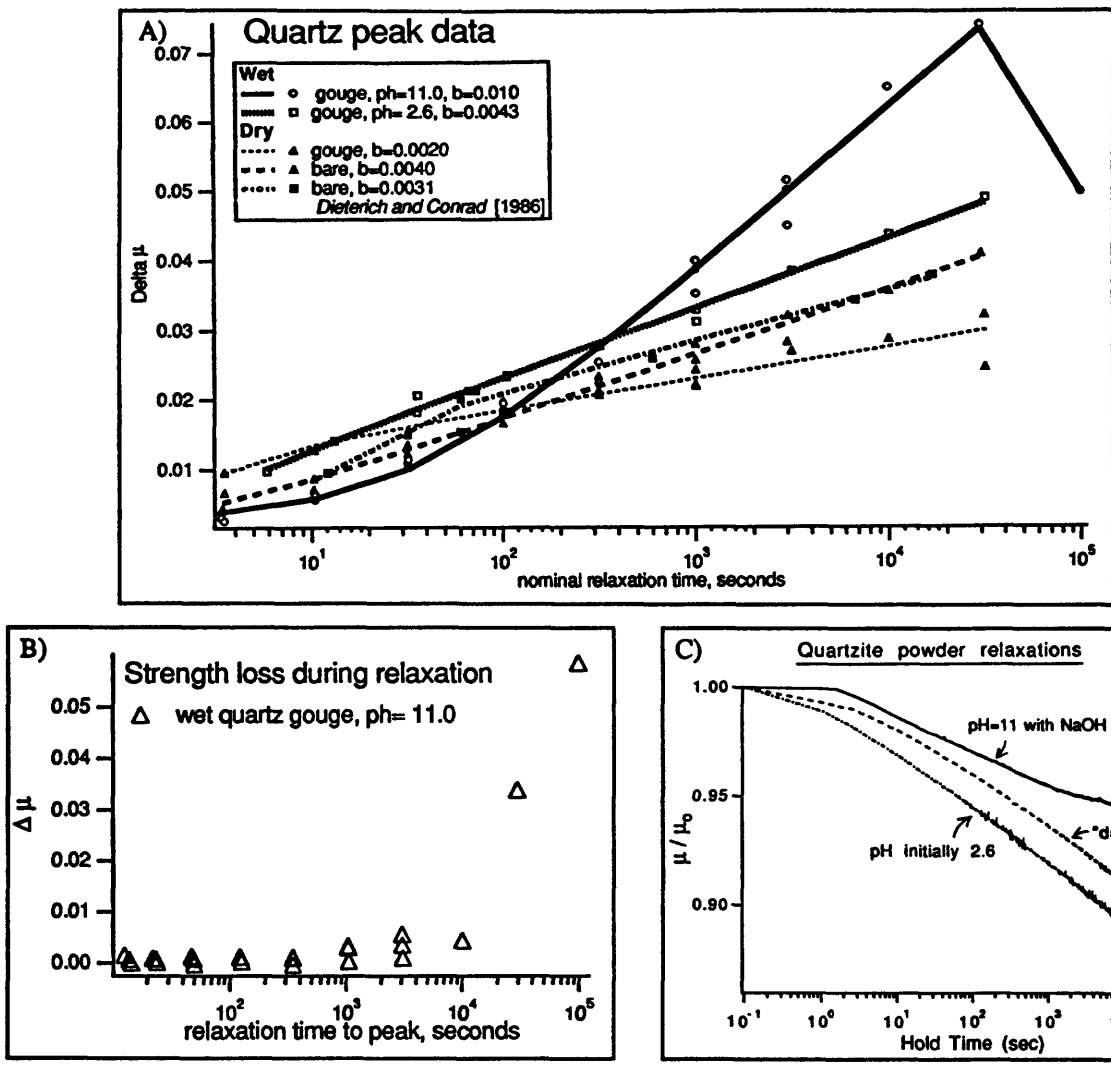



Figure 2. Data from slide-hold-slide tests on powdered quartz layers viewed three different ways. These all show different aspects indicating that the solution used in the pH 11 has altered the frictional behavior of the quartz powder. (A) Static strengthening as a function of log(hold time) for several different runs, including the pH 11 run and one from *Dieterich and Conrad* (1984). Note that most of the data fall reasonably well into a single group, but that the slope for the pH 11 run is substantially higher than the rest. In addition, the final point is considerably lower, we believe because a new mechanism is becoming important, possibly pressure solution. Delta μ is the difference between the transient strength peak following the hold minus the steady-state level following the hold. The level following was used because of offsets in strength after the two longest holds, as shown in (B). (B) Strength loss during hold on quartz powder saturated with pH 11 fluid. The steady-state frictional strength reached after the two longest holds was substantially lower, as shown by this plot of the steady-state level before, minus the steady-state level after. (C) Coefficient of friction versus log(time) showing the relaxation during the longest holds from three runs on quartz, the room dry run, and the two run with altered pH. The nearly straight relaxations seen for the dry and pH 2.6 runs are the expected behavior when an evolution effect is active. The sharp drop-off at the end of the relaxation at pH 11 indicates that a different mechanism, probably a flow-type mechanism, was active at the lowest velocities. This type of behavior is seen in quartz under hydrothermal conditions and is due to pressure solution (*Chester and Higgs*, 1992) and also in serpentine at room temperature (*Reinen*, 1993), in which the new mechanism is probably dislocation glide.

# Fluid-Rock Interaction and Weakening in Faults of the San Andreas System: Inferences from San Gabriel Fault-Rock Geochemistry and Microstructures

James P. Evans
Department of Geology
Utah State University
Logan, UT 84322-4505

and
Frederick M. Chester
Department of Earth and Atmospheric Sciences
Saint Louis University
3507 Laclede Ave.
St. Louis, MO 63103

**Abstract** We examine the influence of fluids on the deformation mechanisms in narrow strike-slip faults of the San Andreas system by examining microstructures and whole rock geochemistry of fault-related rocks. Microstructures of fine-grained ultracataclasites which form the 5-15 cm core of the San Gabriel fault developed by milling and particulate flow of clay-zeolite mixtures. The fault core evolved from the 30-100 m thick cataclasite and damaged zone. Cataclasites adjacent to the ultracataclasite contain numerous cross-cutting and deformed zeolite veins, which indicate repeated fluid flow and deformation cycles during fault development. Whole-rock geochemical analyses suggest that up to 60-80% of rock volume was lost in the fault core during slip on the fault. Volume loss was accomplished by hydrolysis of feldspar and removal of $SiO_2$ and $Al_2O_3$. Fluid-rock interaction and volume loss is a function of net slip, as calculated volume loss is less for a fault splay with less displacement. Fluid-rock interactions recorded along the eastern San Gabriel fault differ markedly from that recorded to the west. This suggests that the fault was hydrologically segmented during its history. The fluid flow properties of the surrounding rocks and internal fault zone properties may vary along the length of a fault to create a hydrologically, and therefore structurally, segmented fault.

***Introduction*** A wide array of geological and geophysical research shows the San Andreas fault is a weak fault within a strong bounding crust. Heat flow measurements (Brune, et al., 1969; Lachenbruch and Sass, 1980), in-situ stress measurements (Zoback and Healy, 1992; Hickman, 1991) and focal mechanism data (Jones, 1988) all show that along much of the fault the maximum principal stress is at a high angle to the fault plane. Mechanisms to explain this behavior include high pore-fluid pressures in the fault, due either to fluids migrating to and through the fault during the seismic cycle (Rice, 1992; Chester et al., 1993), or due to fluids trapped in fine-grained gouge zones (Byerlee, 1990). Non-frictional, possibly diffusive mass transfer processes which have heretofore not been documented along the San Andreas fault, may also dominate along the fault, or some weakening process induced by dynamic slip (Heaton, 1990; Hickman, 1991; Brune,

1991) may cause the weak fault-strong crust behavior of the San Andreas fault system.

Like most seismogenic faults, the San Andreas fault is can be divided into rupture segments based on neotectonic investigations (Schwartz and Coppersmith, 1984) and focal mechanism data ( Jones, 1988). The stress state near the fault varies along the strike of the fault, with different segments subjected to different orientations or relative magnitudes of principal stresses. Coincidence between rupture segments and the present-day state of stress variation on the fault suggests that there is a strong correlation between the stress state and the rupture processes (Jones, 1988).

Recent work by Chester et al. (1993) on the San Gabriel fault and other ancient faults of the San Andreas system suggests that slip can be localized along zones 10's of cm wide through a process of reaction weakening and sub-microscopic particulate flow. Fluid-rock reactions result in abundant, localized alteration to ultra-fine grained clay and zeolite mixtures, which may become entrained as very fine-grained matrix in fault cores (Chester et al., 1993). Data presented by Chester et al. (1993) suggest that fluids penetrated the San Gabriel fault zone throughout its slip history and episodic flow may have been related to rupture events.

We follow up on the work of Chester et al. (1993) and investigate the structure and geochemistry of rocks from the San Gabriel fault to determine how fluids interacted with the fault-related rocks at depth. We present the results of scanning electron microscopy which determine the microstructures of the fault-related rocks and the nature of fluid-rock interactions. Whole-rock geochemistry of fault-related rocks quantifies the degree of fluid-rock interaction across the fault, and provides estimates of the volume loss associated with slip along the San Gabriel fault. We briefly discuss the along-strike variations in fault geochemistry and structure and their implications for segmented structure of the fault. Since the San Gabriel fault is interpreted to have been the active trace of the San Andreas system in the past (Crowell, 1981; Ehlig, 1981; Powell and Weldon, 1992; Powell, 1993), these results have significant implications for understanding the stress state and mechanisms along this large active fault.

***Geologic Setting*** The San Gabriel fault is a component of the San Andreas system in the central Transverse Ranges of southern California (Crowell, 1981; Ehlig, 1981; Powell and Weldon, 1992). Slip on the fault between ~ 12 Ma and 5 Ma accounted for 42-45 km of right-lateral separation (Powell and Weldon, 1992; Powell, 1993). Uplift of the San Gabriel Mountains since the Pliocene due to thrusting and regional uplift exposed crustal levels which were as deep as 2-5 km during slip on the San Gabriel fault (Oakshott, 1971). Microstructures and mineral assemblages of the fault rocks from the San Gabriel fault are consistent with faulting at several kilometers depth (Chester et al., 1993; Anderson et al., 1983). In the central San Gabriel Mountains, the San Gabriel fault diverges into the northern and southern branches, and the southern branch joins the frontal Cucamonga-Sierra Madre thrust. Our investigation examines the north branch of the San Gabriel fault between the junction with the South Branch to the west and the San Antonio fault to the east (Fig. 1). This portion of the North Branch and surrounding basement terrane has been mapped mostly in reconnaissance (Ehlig, 1975; Crowder, 1967; Oakeshott, 1958; Carter and Silver, 1972; Evans, 1982; Jennings and Strand, 1969; Rodgers, 1967). Locally intense deformation associated with the North Branch can be distinguished from general deformation of the basement terrane (e.g.,

Crowder,1967; Ehlig, 1973). On the basis of offset contacts in the basement complex, the principal fault of the North Branch investigated here accommodated approximately 16 km right-lateral slip. Slip on the north branch of the San Gabriel fault appears to diminish eastward. Our sample localities are on the main strand of the North Branch, and on a splay fault with ~6-8 km of slip. The North Branch of the San Gabriel fault may be one of the most deeply and well-exposed traces of the modern San Andreas fault system, and allows us to examine rocks formed at depth along the fault.

Principal fault surfaces of the North Branch juxtapose contrasting rock types and contain a continuous gouge or ultracataclasite layer within an intensely brecciated zone (Chester et al., 1993). The central ultracataclasite layer is 2-10 cm thick and is distinguished by extremely fine grain size and lack of porphyroclasts (Chester et al., 1993; Anderson et al., 1983; Ehlig, 1975). The ultracataclasite layer is bounded by a meters-thick zone of foliated cataclasite, which is bounded by a zone 30 to 100 meters thick of damaged host rock containing cataclasite, fractures and subsidiary faults (Chester et al., 1993). The ultracataclasite layer and foliated cataclasite is referred to as the fault core. The degree of alteration to zeolites and clays increases near the boundary between damaged zone and fault core. Cross-cutting and variably deformed veins and clays indicate alteration and mineralization occurred episodically during fault evolution (Chester et al., 1993). The fault core is largely comprised of fine grained illite/smectite, montmorillonite, and laumontite. The mesoscopic structure of the fault zone records extreme localization of slip to the fault core.

*Microstructures of the fault-related rocks* Optical microscopy and X-ray diffraction studies of samples from three localities (Fig. 1) enable us to identify mineralogy and deformation mechanisms across the fault zone (see also Chester et al. 1993). Deformation in the cataclasite zones involved brittle fracture and incipient cataclasis. Protolith quartz and feldspar grains exhibit intra- and intergranular fracture, and small degrees of alteration and neominerailzation is observed (Chester et al., 1993). As the fault core is approached, deformation intensity (measured by the fracture density) and mineralogic alteration increases dramatically. Near the boundary of the fault core, laumontite veins and illite/smectite increases. Several generations of laumontite veins are evident; we see cross-cutting relationships (Fig. 2a) in which folded and crumpled veins are cut by later, less-deformed veins. Veins are interlayered with clay layers, and the mixtures display folds and irregular "rollups". These observations indicate that: 1) the veins formed episodically over the history of the samples, 2) the veins were emplaced parallel to clay layers in some cases, and 3) the clays developed before and during slip along the fault.

Samples from the central ultracataclasite core exhibit few clues at the optical scale regarding the mechanisms by which most of the strain was accommodated. The core is comprised of very fined grained illite/smectite, montmorillonite, and laumontite. Lamontite is difficult to see in thin sections, occurring as small rounded blebs where visible. Clay-rich samples from the edges of the ultracataclasite in places have two sets of shear bands which are at 20 - 30° to the dominant shear plane of the fault with a synthetic sense of slip (Fig. 2 b), which resemble S-foliation planes of an S-C shear zone (Chester and Logan, 1987). These fabrics reflect from simple shear within the fault, where continuous flow was accommodated by the very fine-grained matrix, and the

discontinuous, simple shear component of deformation was accommodated by slip on these shear bands (Chester and Logan, 1987). Other samples of ultracataclasite exhibit more random fabrics (Fig. 2 c).

To help determine deformation mechanisms and chemical variations of the very fine-grained parts of the fault, we examined samples with scanning electron microscopy (SEM) using secondary and back-scattered electron and energy dispersive X-ray analyses to characterize the samples. The SEM images (Fig. 3) show that cataclasis of feldspar and quartz grains gives way to development of very fine-grained, irregular "groundmass" of grains 20 $\mu$m and less. The ground mass is enriched in Fe, Mg, Na, K, Ca, and Ti (Fig. 3 b and c) and contains scattered discrete grains in an amorphous matrix (Fig. 3 a and b). Very high magnification of the matrix reveals rounded, conglobated and irregular masses that have composition of clay. We do not observe well-defined clay platelets observed in other fault zones (Rutter et al., 1986; Evans, 1990). The very fine-grained matrix of the cataclasite and ultracataclasite may be a retrograde, post-deformation alteration to clay, but we have yet to observe pseudomorph textures to support this.

In places, irregularly shaped pods of uniform Z-contrast are composed of distinct grains (Fig. 3 a). We interpret this texture to be the result of development and cementation of grains into the pods. Element maps of these zones suggest they are laumontite, and thus may be remnants of veins which have been reworked in the cataclasite. SEM observations of the ultracataclasite at magnifications ranging from 40 to 3000 x show no planar fabrics (Fig. 3 c). Occasional rounded to sub-angular fragments appear in an amorphous matrix of material composed of Fe, Mg, Ca, Al, and Si (Fig. 3 c).

The composition of the very fine-grained groundmass which comprises most of the ultracataclasite is compatible with a clay-laumontite mixture (Fig. 3c). Most of the shear strain is confined to this narrow layer. The precise deformation mechanisms within this zone are not known. Mechanisms which may have been responsible for these textures include solution-assisted comminution and/or particulate flow of soft grains of clay and laumontite 20 mm and less in length. Along with the chemical effects of fluids in the fault zone, localized high pore fluid pressures within the very fine grained ultracataclasite may have accompanied the comminution or flow. Previous studies of the deformation of clay- and mica-rich rocks (Evans, 1990; Kanaori et al., 1991) indicate that very fine-grained phyllosilicates may become entrained in the fault by interlayer slip or shredding. Textures of the San Gabriel ultracataclasite are similar to those produced in high strain shear experiments (Yund et al., 1990), and the comminution processes which likely produced the amorphous experimental gouge may be analogous to those which operated in the San Gabriel fault. Although solution assisted processes were important in the fault cores, we do not see discrete pressure solution seams of insoluble residue, truncated or interpenetrating grains, or fibrous overgrowths characteristic of diffusive mass transfer processes.

Non-frictional deformation mechanisms in the core of the fault may help resolve part of the fault-strength problem along the San Andreas system (Hickman, 1991). Laboratory studies of frictional properties of laumontite (Morrow and Byerlee, 1989) and clays (Moore et al., 1989) all suggest that the frictional strength of these minerals exceeds that predicted in most weakening models of faults (e. g., Rice, 1992). However,

if a mixture of these minerals deforms fluid-enhanced comminution and/or particulate flow of micron-sized particles in the presence of fluid, then low-strain, short-term, isochemical frictional experiments may not strictly apply to the natural processes, and high-strain experiments (Yund et al., 1990) with chemical interactions between fluids and deforming materials (Blanpied et al., 1992) may reveal the processes which were active in the fault zones.

***Fault rock geochemistry*** Whole-rock geochemical analyses of samples from the fault zones quantify chemical variations associated with deformation and can be used to estimate volume losses associated with deformation (O'Hara, 1988). Geochemical analyses of fault-related rocks must be compared to analyses of undeformed rock in order to determine the chemical changes associated with deformation. Much of the exposed part of the North Branch of the San Gabriel fault occurs in granodiorite, but locally the fault cuts diorite, biotite gneiss, and rarely cuts small amphibolite pods. Because the fault experienced up to 16 km of displacement, rocks in the fault zone may have interacted mechanically and chemically with different protoliths. Thus, we use an average, representative protolith composition determined by averaging the results of geochemical analyses of many different samples of the protolith (excluding amphibolite) from each site examined. We exclude the amphibolite analysis because of its volumetric unimportance in the San Gabriel Range.

X-Ray fluorescence analysis of major, minor, and trace elements performed on 36 samples (Table 1) from three sites indicate significant loss of $SiO_2$ and $Al_2O_3$, and minor losses of $K_2O$ and NaO in the fault zones as compared with representative protoliths (Fig. 4a). Comparison of fault rocks with protolith compositions suggest little or no mobility of Ti, Fe, Mg, Mn, P, and Ca. We use Ti (as $TiO_2$) as the immobile constituent in our calculations of volume changes within the fault. Geochemical studies of metasomatized rocks and shear zones suggest the immobility of Ti as well as Mg, Mn, Zr, Y, and P (Correns 1978, Dostal *et al.* 1980, Floyd & Winchester 1983, Winchester & Max 1984, Sinha *et al.* 1986, O'Hara 1988, O'Hara & Blackburn 1989, Glazner & Bartley 1991). The enrichment of Ti and associated high-field strength cations is due to volume losses of $SiO_2$ and alkalies.

Rocks from the ultracataclasite and adjacent foliated cataclasite in the fault core display the largest enrichment of immobile elements and depletion of $SiO_2$ (Figs. 4 and 5). The enrichment of immobile elements is interpreted to be due to the preferential loss of mobile elements, especially Si (Fig. 4). The loss of $SiO_2$ increases with slip and cataclasis. Samples from the splay fault at the North Fork site exhibit less $SiO_2$ depletion than the other localities, and the loss of $SiO_2$ is greatest in the ultracataclasite core of the faults. The chemical changes observed in the foliated cataclasite appear to be as great as in the ultracataclasite, which suggests the ultracataclasite evolved geochemically from the foliated cataclasite. The variations in chemistry observed here are the result primarily of hydrolysis reactions which convert feldspars to clays and to the development of zeolites. The reactions which appear to have occurred in these fault-related rocks include (Garrels and Christ, 1965; Faure, 1991):

$$\underset{\text{albite}}{2NaAlSi_3O_8} + 9\,H_2O + 2H^+ \leftrightarrow \underset{\text{kaolinite}}{Al_4Si_4O_{10}(OH)_8} + 2Na^+ + \underset{\text{dissolved silica}}{8H_4SiO_4}$$

$$2NaAlSi_3O_8 + Ca^{2+} + 4H_2O \leftrightarrow CaAl_2Si_4O_{12} \cdot 4H_2O + 2Na^+ + 2SiO_2$$
albite laumontite (amorph)

$$SiO_2 + H_2O \leftrightarrow H_4SiO_4$$
quartz dissolved silica

$$7Al_2Si_2O_5(OH)_4 + 8SiO_2 + Mg^{2+} \leftrightarrow 6[(Al_2)(Si_{3.67},Al_{0.33})O_{10}(OH)_2]Mg^{2+}_{0.167} + 7H_2O + 2H^+$$
kaolinite Mg-montmorillonite

We estimate the relative amount of volume loss and fluid-rock interaction by examining the changes in element abundances relative to the immobile element(s). To prevent a single whole rock analysis from skewing the results, we examine the variations of calculated oxides for mean values of samples from the damaged zone, the cataclasite, and the ultracataclasite core of the fault. A simple way to determine the volume loss in the fault zone is to examine the slopes in Fig. 5. The slope of the immobile isocon (Grant, 1986) reveals the mass change, and, if the density of the protolith and fault rocks are known, the volume change in the alteration. An immobile isocon slope of one indicates isovolumetric conditions, that is, volumetric concentrations remained constant throughout deformation. Cataclasites and ultracataclasites from two sites along the North Branch of the San Gabriel fault yield slopes of 3-5, or volume losses relative to the protolith of 60-80% for the ultracataclasites, 50-75% for cataclasites, and about 30% volume loss for rocks from the North Fork site, a subsidiary fault with smaller displacement.

Volume factors (Fig. 5) determined using the method of Grant (1986) can be combined with Gresen's (1967) formulations to write reactions responsible for the alteration of protolith to fault rock during deformation. For example, the reaction for the ultracataclasite at Bear Creek, relative to the protolith, can be written as:

100 g protolith - 51.52 g $SiO_2$ - 8.49g $Al_2O_3$ - 2.38 g$K_2O$ -3.05g $Na_2O$ -0.236g $P_2O_5$- 1.32g CaO-0.07g MnO = .129 g $TiO_2$ +0.29 g MgO +0.483g $Fe_2O_3$+ 34 g fault rock

The fluid interaction with the fault-rocks resulted in development of a clay-zeolite fault core, volume loss, and intense strain localization on the fault. Volume loss from the water-rock interaction may have accompanied compaction of the fault during the seismic cycle (Sleep and Blanpied, 1992), and fluid flow may have been episodic due to variations in fault zone permeability at different points in the seismic cycle (Chester et al., 1993; Sibson, 1992).

***Discussion*** The results here are markedly different from the most results of work on the western portion of the San Gabriel fault (Anderson et al., 1983), who showed limited water-rock interaction based on little clay or zeolite content. Several localized regions along the western San Gabriel fault exhibit similar textures to what is described above, but overall Anderson et al. (1983) suggested that the western portion of the fault developed under relatively dry conditions. Our work indicates significant amounts of fluid-rock interaction during fault zone evolution. Why the difference?

We suggest that the contrasting results may demonstrate a hydrologic

segmentation or partitioning of the fault. Different parts of the fault may have been subjected to different hydrologic regimes, which in turn resulted in very different fault-zone hydrology processes. Topographic relief, different country rock types, variations in regional permeability structure, or local variation in climate may all conspire to provide more or less water to different parts of a long fault (Forster and Evans, 1991). Laumontite mineralization may mark fluid flow in and near parts of faults of the San Andreas system (James and Silver, 1988; Chester et al., 1993), whereas other portions of faults are marked by little zeolite development (Anderson et al., 1983). The different degree of available water may result in varying fluid pressure regimes and different amounts of chemical alteration along the length of the fault. Some portions of the fault may experience variation of fluid pressure during the seismic cycle (Sibson, 1992); some parts of the fault may remain relatively dry during its history. Combinations of fluid pressurization and chemical alteration results in weak faults (Chester et al., 1993 Sibson, 1992; Rice, 1992; Byerlee, 1990; Janecke and Evans, 1988). Thus, the mechanical behavior of large segments of faults is at least in part controlled by local and regional hydrogeologic setting of the faults.

In addition to the external hydrogeologic parameters which may cause fluid partitioning along a fault, internal variation of fault zone structure (Byerlee, 1993) may cause variations in fluid-rock interactions in faults. Pods of high pore-fluid pressure may exist over a variety of scales (Byerlee, 1993), and these variations may be preserved in the rock record by spatial variations of the geochemistry of fault-related rocks.

***Implications for stress state in the fault*** Several explanations for the weak behavior of the San Andreas fault are based on the premise that the mean stress and the least principal stress ($\sigma_3$) are greater in the fault core than in the surrounding host rock due to presence of weaker fault zone materials (Rice, 1992; Chester et al., 1993; Byerlee, 1990). This stress state could allow pore-fluid pressures in the core to exceed $\sigma_3$ in the host rock. There may be a mechanical problem with this explanation in that high pore fluid pressure could exist at the boundary between the fault core and host rock, which would induce hydrofracture in the host rock and thereby prevent the build-up of high pore pressure (Scholz, 1992). However, the fault core to host-rock transition is gradational (Chester et al., 1993) so both $\sigma_3$ and pore fluid pressure may increase gradually towards the ultracataclasite layer and locally $\sigma_3$ would nearly always be greater than the pore pressure. If during some portion of the seismic cycle the pore pressure exceeds $\sigma_3$ locally, limited amounts of hydrofracture may occur. These excursions in pore pressure may be the cause of the veining observed in the San Gabriel fault (Chester et al., 1993).

***Conclusions*** We show that the very fine-grained ultracataclasite core of part of the San Andreas fault developed by fluid-assisted comminution, alteration, and particulate flow. Episodic veining documents periodic increases in pore fluid pressure in the fault. Cataclasis, fluid-assisted alteration, mineralization, and fault-normal compaction operated episodically during fault evolution, and each process probably operated at its own characteristic time scale. The chemical fluid-rock interaction resulted in significant alteration of feldspar, dissolution and transport of $SiO_2$ and to a lesser degree $Al_2O_3$. Syndeformation reactions resulted in volume loss across the fault, which accommodated compaction of the core into a narrow, high shear strain zone. Geochemical data

presented here combined with microstructures, mineralogic data, and stress inversions of small faults adjacent to the San Gabriel fault (Chester et al., 1993) show the narrow ultracataclasite fault core and cataclasite bounding zone was oriented at a high angle to the local maximum principal stress and was the locus of intense syn-deformation fluid-rock interaction. The combination of these processes during the history of the fault may explain the weak behavior of faults of the San Andreas system. Along-strike variations in the magnitudes of fluid pressure and fluid-rock interaction my result in along-strike variations in the rheology of fault segments. This rheologic segmentation may be reflected by different stress states and rupture patterns along the faults.

*Acknowledgements*: Funding for this work was provided by U. S. Geological Survey NEHRP grant 1434-92-G-2184 and National Science Foundation Grants EAR-9205774 and EAR-9205973.

**References**

Anderson, J. L., Osbourne, R. H., and Palmer, D. F. 1983, Cataclastic rocks of the San Gabriel fault-An expression of deformation at deeper crustal levels in the San Andreas fault, Tectonophysics, v. 98, p. 209-251.

Blanpied, M. L., Lockner, D. A., and Byerlee, J. D., 1992, An earthquake mechanism based on rapid sealing of faults, Nature, v. 358, p. 574-576.

Brune, J. N., 1991, Seismic source dynamics, radiation and stress, Reviews of Geophysics, supplement to v. 29, p. 688-699.

Brune, J. N., Henyey, T. L., and Roy, R. F., 1969, Heat flow, stress and rate of slip along the San Andreas fault, California, Journal of Geophysical Research, v.74, p. 3821-3827.

Byerlee, J., 1993, Model for episodic flow of high-pressure water in fault zones before earthquakes, Geology, v. 21, p. 303-306.

_____, 1990, Friction, overpressure, and fault normal compression, Geophysical Research Letters, v. 17, p. 2109-2112.

Carter, B., and Silver, L. T., 1972, Structure and petrology of the San Gabriel anorthosite-syenite body, California, Int. Geol. Cong., Proceedings, section 2, p. 303-311.

Chester, F. M., Evans, J. P., and Biegel, R. L., 1993, Internal structure and weakening mechanisms of the San Andreas Fault, Journal of Geophysical Research, 98, 771-786.

Chester, F. M., and Logan, J. M., 1987, Composite planar fabric of gouge from the Punchbowl fault, California, Journal of Structural Geology, v. 9, p. 621-634.

_____, 1986, Implications for mechanical properties of brittle faults from observations of the Punchbowl fault zone, California, Pure and Applied Geophysics, v. 124, p. 80-106.

Correns, C. W., 1978, Titanium: behavior in metamorphic reactions. In: Handbook of Geochemistry (edited by Wedepohl, K. H.). Springer, Berlin.

Crowder, D. F., 1967, Mineral resources of the Devils Canyon-Bear Canyon primitive area, California, U. S. Geological Survey Bulletin, 1230-G, 21 p.

Crowell, J. C., 1981, An outline of the tectonic history of southeastern California, in: Ernst, W. G., ed., The geotectonic development of California, (Rubey Volume 1), Prentice-Hall, Englewood, New Jersey, p. 584-600.

Dostal, J., Strong, D. F., and Jamieson, R. A., 1980, Trace element mobility in the mylonite zone within the ophiolite aureole, St. Anthony complex, Newfoundland, Earth & Planetary Science Letters, v. 49, p. 188-192.

Ehlig, P. L., 1973, History, seismicity, and engineering geology of the San Gabriel fault, in: Geology, Seismicity and Environmental Impact, Spec. Pub. Assoc. Engineering Geol., 247-251.

_____, 1975, Basement rocks of the San Gabriel Mountains, south of the San Andreas fault in southern California, Calif. Div. Mines and Geol. Sp. Report 118, p. 177-186.

_____, 1981, Origin and tectonic history of the basement terrain of the San Gabriel Mountains, central Transverse Ranges, in: in: Ernst, W. G., ed., The geotectonic development of California, (Rubey Volume 1), Prentice-Hall, Englewod, New Jersey, p. 584-600.

Evans, J. G., 1982, The Vincent thrust, eastern San Gabriel Mountains, California, U. S. Geological Survey Bulletin 1507.

Evans, J. P., 1990, Textures, deformation mechanisms, and the role of fluids in the cataclastic deformation of granitic rocks. In: Deformation mechanisms, rheology and tectonics (edited by Knipe, R. J. & Rutter, E. H.), Geol. Soc. London Special Pub. 54, 29-39.

_____, 1988, Deformation mechanisms in granitic rocks at shallow crustal levels, Journal of Structural Geology, v. 10, p. 437-443.

Faure, G., 1991, Principles and Applications of Inorganic Geochemistry, Macmillan Publishing Company, 250-309.

Floyd, P. A., and Winchester, J. A., 1983, Element mobility associated with meta-shear zones within the Ben Hope amphibolite suite, Scotland, Chemical Geology v. 39, p. 1-15.

Forster, C. B., and Evans, J. P., 1991, Hydrogeology of thrust faults and crystalline thrust sheets: Results of combined field and modeling studies, Geophysical Research letters, v. 18, p. 979-982.

Garrels, R. M., and Christ, C. L., 1965, Solutions, Mineral, and Equilibria, Freeman, Cooper & Company, 438 p.

Glazner, A. F., and Bartley, J. M., 1991, Volume loss, fluid flow and state of strain in extensional mylonites from the central Mojave Desert, California, Journal of Structural Geology, v. 13, p. 587-594.

Grant, J. A., 1986, The isocon diagram-A simple solution to Gresens' equation for metasomatic alteration, Economic Geology, v., 81, p. 1976-1982.

Gresens, R. L., 1967, Composition-volume relationships of metasomatism, Chemical Geology v. 2, 4 p.7-65.

Heaton, T. H., 1990, Evidence for and implication of self-healing pulses of slip in earthquake rupture, Physics Earth Planetary Interiors, v. 64, p. 1-20.

Hickman, S. H., 1991, Stress in the lithosphere and the strength of active faults, U.S. Nat. Rep. to Internat. Un. of Geodesy and Geophys. 1987-1990, Reviews of Geophys., Suppl., 759-775.

James, E. W., and Silver, L. T., 1988, Implications of zeolites and their zonation in the Cajon Pass deep drillhole, Geophysical Research Letters, v. 15, p. 973-976.

Janecke, S. U., and Evans, J. P., 1988, Feldspar-influenced rock rheologies, Geology, v. 16, p. 1064-1067.

Jennings, C. W., and Strand, R. G., 1969, Los Angeles Sheet: Geol. Map of California, Calif. Div. Mines Geol., scale 1:250,000.

Jones, L. M, 1988, Focal mechanisms and the state of stress on the San Andreas fault in southern California, Journal of Geophysical Research, v. 93, 11,883-11,844.

Kanaori, Y., Kawakami, S., and Yairi, K, 1991, Microstructure of deformed biotite defining foliation in cataclastic zones in granite, central Japan, Journal of Structural Geology, v. 11, p. 777-786.

Lachenbruch, A. H., and Sass, J. H., 1980, Heat flow and energetics of the San Andreas fault zone, Journal of Geophysical Research, v. 85, p. 6185-6222.

Mitra, G., 1984, Brittle to ductile transition due to large strains along the White Rock thrust, Wind River Mountains, Wyoming, Journal of Structural Geology, v. 6, p. 51-61.

Moore, D. E., Summers, R., and Byerlee, J. D., 1989, Sliding behavior and deformation of heated illite gouge, Journal of Structural Geology, v. 11, p. 329- 342.

Morrow, C. A., and Byerlee, J. D., 1991, A note on the frictional strength of laumontite from Cajon Pass, California, Geophysical Research Letters, v. 18, p. 211-214.

O'Hara, K., 1988, Fluid flow and volume loss during mylonitization: An origin for phyllonite in an overthrust setting, North Carolina, U.S.A. Tectonophysics v. 156, p. 21-36.

O'Hara, K., and Blackburn, W. H., 1989, Volume loss model for trace-element enrichment in mylonites, Geology, v. 17, p. 524-527.

Oakeshott, G. B., 1958, Geology and mineral deposits of San Fernando quadrangle, Los Angeles County, California, Calif. Div. Mines and Geol. Bull. 172, 147 p.

_____, 1971, Geology of the epicentral area, Calif. Div. Mines Geol., v. 196, p. 19-30.

Powell, R. E., 1993, Balanced palinspastic reconstruction of Pre-Late Cenozoic Paleogeology, Southern California: Geologic and kinematic constraints on evolution of the San Andreas fault system, in: Powell, R. E., Weldon, R. J., and Matti, J. C., eds., The San Andreas Fault System: Displacement, Palinspastic Reconstruction, and Geologic Evolution, Geol. Soc. Am. Memoir 178, p. 1-107.

Powell, R. E., and Weldon, R. J., 1992, Evolution of the San Andreas Fault, Ann. Rev. Earth and Plan. Sci. 20, 431-468.

Rice, J. R., 1992, Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault, in: Fault Mechanics and Transport Properties in Rocks (eds. Evans, B. & Wong, T. F. Academic Press, p. 475-503.

Rodgers, T. H., compiler, 1967, San Bernardino Sheet: Geol. Map of California, Calif. Div. Mines Geol., scale 1:250,000.

Rutter, E. H., Maddock, R. H., Hall, S. H., and White, S. H., 1986, Comparative microstructures of natural and experimentally produced clay-bearing fault gouges, Pure and Applied Geophysics, v. 124, p. 3-30.

Scholz, C. H., 1992, Earthquakes-weakness amidst strength, Nature, v. 359, p.677-678.

Schwartz, D. P., and Coppersmith, K. J., 1984, Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, Journal of Geophysical Research, v. 89, p. 5681-5698.

Sibson, R. H., 1992, Implications of fault-valve behavior for rupture nucleation, Tectonophysics, v. 211, p. 283-293.

Sinha, A. K., Hewitt, D. A., and Rimstidt, J. D., 1986, Fluid interaction and element mobility in the development of ultramylonites, Geology, v. 14, p. 833-886.

Sleep, N. H., and Blanpied, M. L., 1992, Creep, compaction, and the weak rheology of major faults, Nature, v. 359, p. 687-692.

Winchester, J. A., and Max, M. D., 1984, Element mobility associated with syn-metamorphic shear zones near Scotchport, NW Mayo, Ireland, Journal of Metamorphic Geology, v. 2, p. 1-11.

Yund, R. A., Blanpied, M. L., Tullis, T. E., and Weeks, J. D., 1990, Amorphous material in high strain experimental fault gouges, Journal of Geophsyical Research, v. 95, p. 15,589-15,602.

Zoback, M. D., and Healy, J. H., 1992, In situ stress measurements to 3.5 km depth in the Cajon Pass scientific borehole: Implications for the mechanics of crustal faulting, Journal of Geophysical Research, v. 97, p. 5039-5057.

| Sample | type | $SiO_2$ | $Al_2O_3$ | CaO | MgO | $Na_2O$ | $K_2O$ | $Fe_2O_3$ | MnO | $TiO_2$ | $P_2O_5$ | LOI | Sum | Cr | Rb | Sr | Y | Zr | Nb | Ba |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Devil's Canyon** | | | | | | | | | | | | | | | | | | | | |
| SG1 | gneiss | 50.0 | 5.39 | 16.0 | 15.1 | 0.62 | 0.22 | 8.53 | 0.17 | 0.493 | 0.04 | 3.50 | 100.3 | 1190 | 17 | 64 | <10 | 19 | 25 | <10 |
| SG3 | pc | 50.7 | 15.8 | 5.51 | 3.77 | 3.17 | 1.63 | 7.06 | 0.11 | 1.24 | 0.25 | 10.3 | 99.7 | 85 | 69 | 417 | 41 | 223 | 26 | 332 |
| SG 5 | gne/pc | 61.4 | 14.4 | 3.08 | 2.43 | 3.41 | 1.96 | 6.09 | 0.10 | 0.840 | 0.20 | 5.90 | 100.0 | 52 | 63 | 585 | <10 | 263 | 42 | 832 |
| SG6B | cata/uc | 57.6 | 15.1 | 3.96 | 2.30 | 3.07 | 2.73 | 5.51 | 0.11 | 0.661 | 0.22 | 8.25 | 99.9 | 40 | 92 | 1690 | 38 | 167 | 30 | 1630 |
| SG 7 | gne/pc | 73.7 | 11.8 | 1.35 | 0.96 | 3.86 | 2.23 | 2.42 | 0.22 | 0.222 | 0.04 | 2.65 | 99.6 | < 10 | 61 | 133 | 102 | 295 | 46 | 935 |
| SG 13O | grano host | 73.7 | 14.7 | 1.54 | 0.16 | 4.71 | 4.09 | 0.43 | 0.03 | 0.061 | 0.02 | 0.90 | 100.6 | < 10 | 106 | 542 | 21 | 49 | 36 | 1530 |
| SG 13F | cata | 48.1 | 21.4 | 10.0 | 1.81 | 1.24 | 1.42 | 3.81 | 0.07 | 0.622 | 0.25 | 11.5 | 100.5 | < 10 | 53 | 1110 | 24 | 228 | 25 | 570 |
| SG 13P | amph host | 44.5 | 12.0 | 11.4 | 13.2 | 1.79 | 0.90 | 11.2 | 0.15 | 1.67 | 0.10 | 2.15 | 99.1 | 581 | 17 | 289 | 30 | 37 | 29 | 201 |
| SG13V | gne/pc | 63.5 | 15.2 | 3.62 | 2.01 | 2.98 | 2.62 | 7.01 | 0.12 | 1.02 | 0.28 | 1.80 | 100.4 | 32 | 125 | 302 | 82 | 433 | 43 | 923 |
| **Bear Canyon** | | | | | | | | | | | | | | | | | | | | |
| SG9A | cata | 49.5 | 17.0 | 6.29 | 3.10 | 2.71 | 1.40 | 6.52 | 0.11 | 0.885 | 0.28 | 11.4 | 99.4 | 18 | 72 | 751 | 30 | 166 | 14 | 971 |
| SG9B | cata | 52.4 | 15.2 | 5.58 | 3.00 | 1.92 | 2.16 | 6.00 | 0.16 | 0.904 | 0.21 | 11.0 | 99.1 | 104 | 86 | 2120 | 26 | 141 | 27 | 2410 |
| SG9C1 | uc | 55.8 | 14.9 | 4.33 | 2.49 | 1.72 | 3.91 | 5.23 | 0.12 | 0.788 | 0.22 | 9.15 | 99.4 | 29 | 63 | 844 | 13 | 112 | 28 | 706 |
| SG9C2 | uc | 50.1 | 16.0 | 9.59 | 2.44 | 1.48 | 1.49 | 5.01 | 0.11 | 0.652 | 0.20 | 13.2 | 100.5 | 79 | 129 | 2660 | 13 | 157 | 39 | 3660 |
| SG9D | cata | 61.6 | 15.0 | 4.18 | 1.47 | 2.76 | 3.32 | 3.28 | 0.06 | 0.472 | 0.12 | 7.85 | 100.3 | 36 | 85 | 570 | <10 | 115 | 33 | 986 |
| SG9E | cata | 52.9 | 16.4 | 6.36 | 3.44 | 2.53 | 1.33 | 6.90 | 0.12 | 0.860 | 0.25 | 9.25 | 100.5 | 14 | 42 | 664 | 23 | 178 | 44 | 658 |
| SG9F1 | cata/pc | 55.2 | 15.7 | 4.87 | 3.33 | 4.20 | 1.66 | 6.42 | 0.12 | 0.811 | 0.24 | 6.30 | 99.1 | 19 | 61 | 617 | 15 | 197 | 33 | 1210 |
| SG9F2 | cata/pc | 50.9 | 17.0 | 6.94 | 3.76 | 3.22 | 1.40 | 7.65 | 0.13 | 0.946 | 0.28 | 7.60 | 100.0 | 20 | 48 | 624 | 25 | 215 | 32 | 774 |
| SG9H | cata/pc | 55.7 | 17.0 | 5.26 | 1.35 | 3.75 | 2.49 | 4.59 | 0.09 | 0.551 | 0.23 | 8.80 | 100.1 | < 10 | 55 | 897 | 42 | 184 | 26 | 1420 |
| SG9I | pc | 64.6 | 15.1 | 2.71 | 1.01 | 4.47 | 2.97 | 3.01 | 0.13 | 0.554 | 0.16 | 5.20 | 100.3 | < 10 | 54 | 632 | 39 | 156 | 31 | 2490 |
| SG9J | pc | 63.7 | 15.4 | 3.50 | 0.73 | 5.72 | 2.36 | 2.17 | 0.07 | 0.333 | 0.08 | 5.40 | 99.8 | <10 | 50 | 744 | 11 | 106 | 22 | 2040 |
| SG991 | uc | 55.2 | 15.4 | 6.20 | 1.89 | 1.39 | 1.70 | 4.07 | 0.10 | 0.554 | 0.15 | 12.7 | 99.6 | 29 | 53 | 817 | 14 | 160 | 20 | 681 |
| SG992 | uc | 62.1 | 14.2 | 4.70 | 1.16 | 2.05 | 2.95 | 2.60 | 0.06 | 0.342 | 0.10 | 10.0 | 100.5 | 29 | 86 | 779 | 12 | 107 | 16 | 1450 |
| SG993 | uc | 56.6 | 14.4 | 4.69 | 1.92 | 1.78 | 2.58 | 4.05 | 0.10 | 0.552 | 0.17 | 12.1 | 99.8 | 42 | 85 | 5360 | 22 | 128 | 32 | 2070 |
| SGC' | uc | 55.2 | 15.2 | 4.27 | 2.59 | 1.72 | 4.22 | 5.49 | 0.13 | 0.739 | 0.21 | 8.05 | 98.5 | 82 | 156 | 2290 | 20 | 167 | 33 | 3550 |
| **North Fork, San Gabriel River** | | | | | | | | | | | | | | | | | | | | |
| SG 22A | mig/pro | 46.5 | 15.1 | 7.19 | 8.45 | 2.68 | 1.68 | 9.45 | 0.15 | 1.61 | 0.31 | 5.60 | 98.9 | 227 | 69 | 556 | 18 | 144 | 41 | 611 |
| SG 22B | pc | 45.5 | 14.7 | 7.24 | 7.45 | 2.38 | 2.21 | 10.1 | 0.17 | 1.92 | 0.34 | 6.50 | 98.7 | 164 | 86 | 572 | 18 | 150 | 48 | 767 |
| SG 22C | pc | 44.9 | 14.9 | 7.57 | 6.41 | 2.66 | 1.04 | 10.6 | 0.18 | 2.11 | 0.39 | 9.00 | 99.9 | 112 | 27 | 584 | 32 | 167 | 44 | 326 |
| SG22 D | cata | 60.9 | 15.5 | 3.24 | 2.63 | 4.66 | 2.98 | 4.63 | 0.10 | 0.621 | 0.17 | 3.95 | 99.6 | 37 | 102 | 467 | 39 | 126 | 33 | 1180 |
| SG 22E | pc/pro | 55.6 | 16.1 | 6.27 | 3.22 | 3.07 | 1.82 | 7.68 | 0.15 | 0.875 | 0.30 | 5.20 | 100.5 | 114 | 107 | 978 | 28 | 237 | 41 | 484 |
| SG 22F | cata | 51.3 | 15.4 | 2.47 | 3.14 | 1.45 | 5.12 | 9.49 | 0.15 | 2.07 | 0.70 | 7.30 | 98.9 | 17 | 207 | 173 | 40 | 330 | 54 | 1690 |
| SG 22H | cata | 57.4 | 16.2 | 2.95 | 3.50 | 3.27 | 3.27 | 6.60 | 0.13 | 0.758 | 0.21 | 5.60 | 100.1 | 55 | 125 | 499 | 22 | 127 | 40 | 1000 |

Table 1. Whole rock geochemical analyses (in weight per cent) of fault-related rocks of the North Branch, San Gabriel Fault, southern California. Cata - cataclasite, uc - ultracataclasite, pc - protocataclasite, pro - protolith, amph - amphibolite.

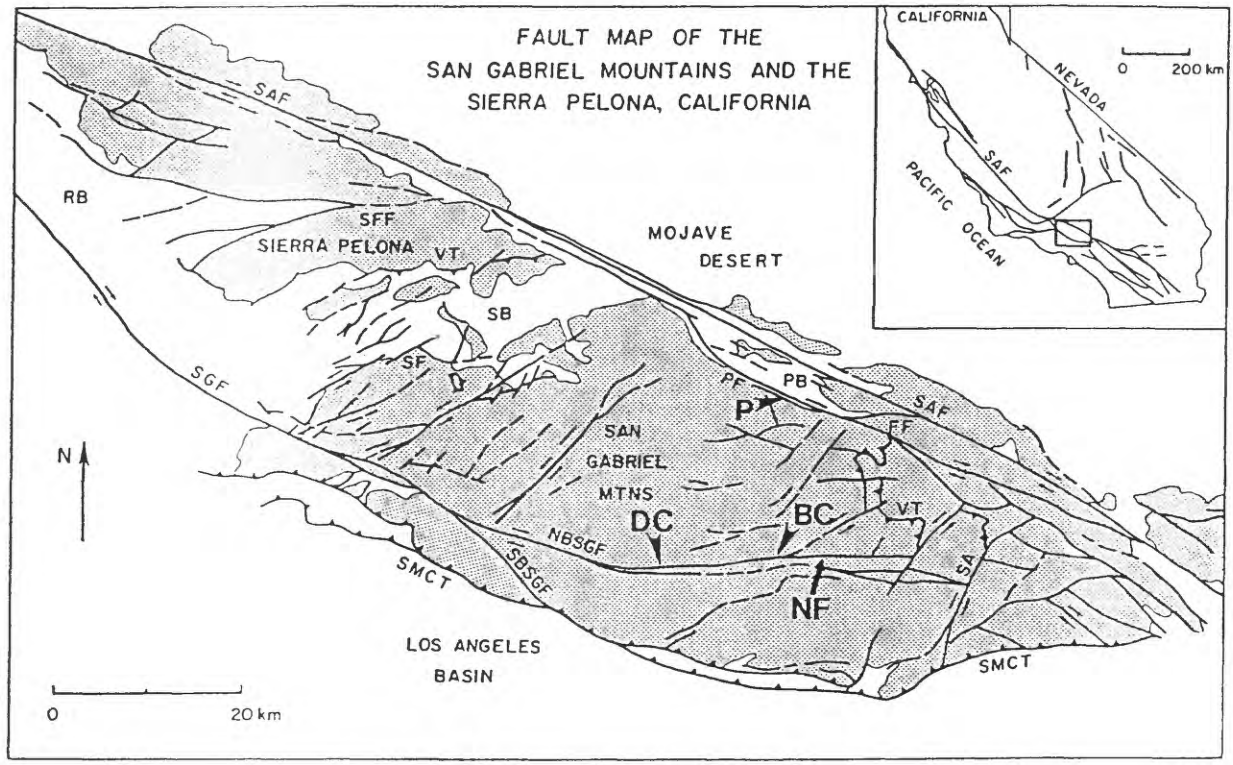



**Figure 1**. Generalized geologic map of the San Gabriel Mountains and vicinity, southern California. Stippled pattern represents crystalline rocks. Bold arrows indicate sampling localities examined in this study. BC - Bear Creek, DC- Devil's Canyon, NF = North Fork site. Other symbols include: SGF - San Gabriel fault (NBSGF-north branch, SBSGF-south branch), SMCT- Sierra Madre-Cucamonga thrust, VT- Vincent thrust, SAF- San Andreas fault, PF - Punchbowl fault, SF - Soledad Fault, SFF - San Francisquito Fault, SA - San Antonio Fault.

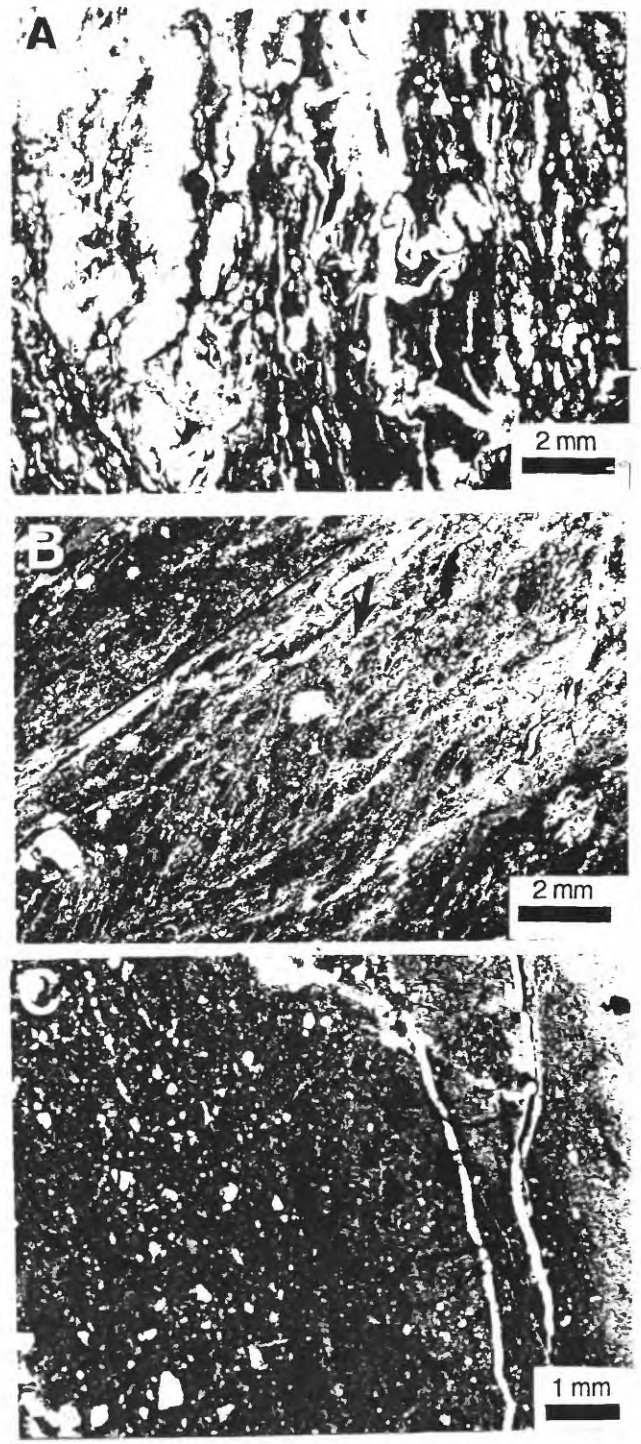



**Figure 2.** Photomicrographs of fault-related rocks, north branch San Gabriel fault.
A. Laumontite veins (white layers) in cataclasites adjacent to the fault core are interlayered with dark clay layers. Some veins are deformed whereas others exhibit a planar fabric. Note cross-cutting veins (arrow) which suggest different episodes of vein development and deformation.
B. Shear-related fabrics at the cataclasite-ultracataclasite boundary. Short arrows point to planar fabrics in clay gouge which form angles of 20-30° to the ultracataclasite boundary (marked by solid line). Fabric indicates right-lateral simple shear along edge of the cataclasite. Fabric is similar to the S-foliation of S-C fabrics, and represent simple shear in clayey gouge (Chester and Logan 1987).
C. Fault core consists of ultracataclasite comprised of quartz fragments in a matrix of often submicroscopic clay and zeolite mixtures.

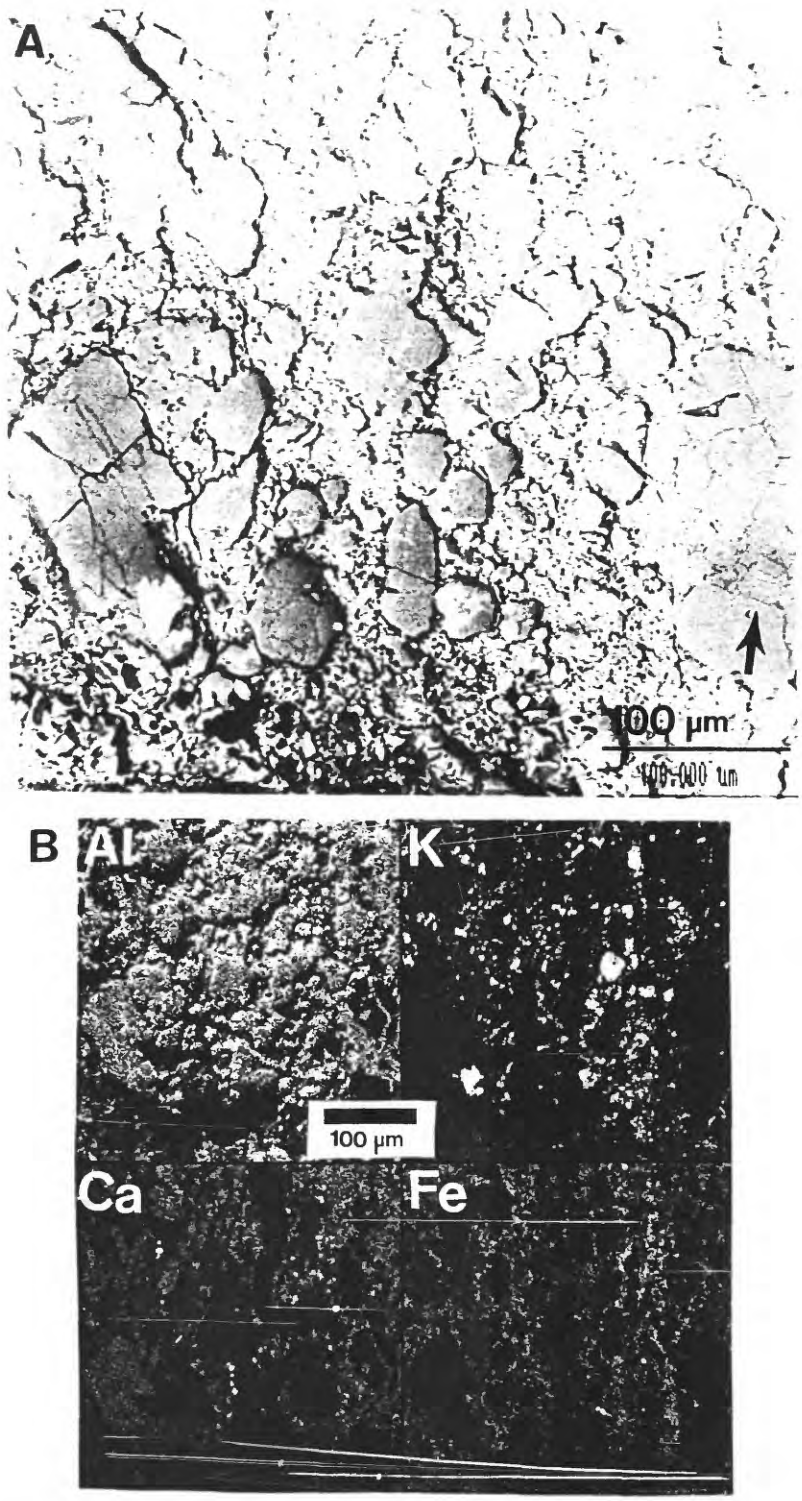



**Figure 3**. Scanning electron images of cataclasite and ultracataclasite.
A. Angular to subrounded fragments of quartz and feldspar lie in fine-grained clay-zeolite matrix. Several "clasts" are actually recemented zeolite, shown by arrow.
B. Element map of image in A. Fine-grained cataclasite is enriched in iron (Fe), and fragments consist primarily of feldspars. Al- Aluminum; K-Potassium; Ca- Calcium.

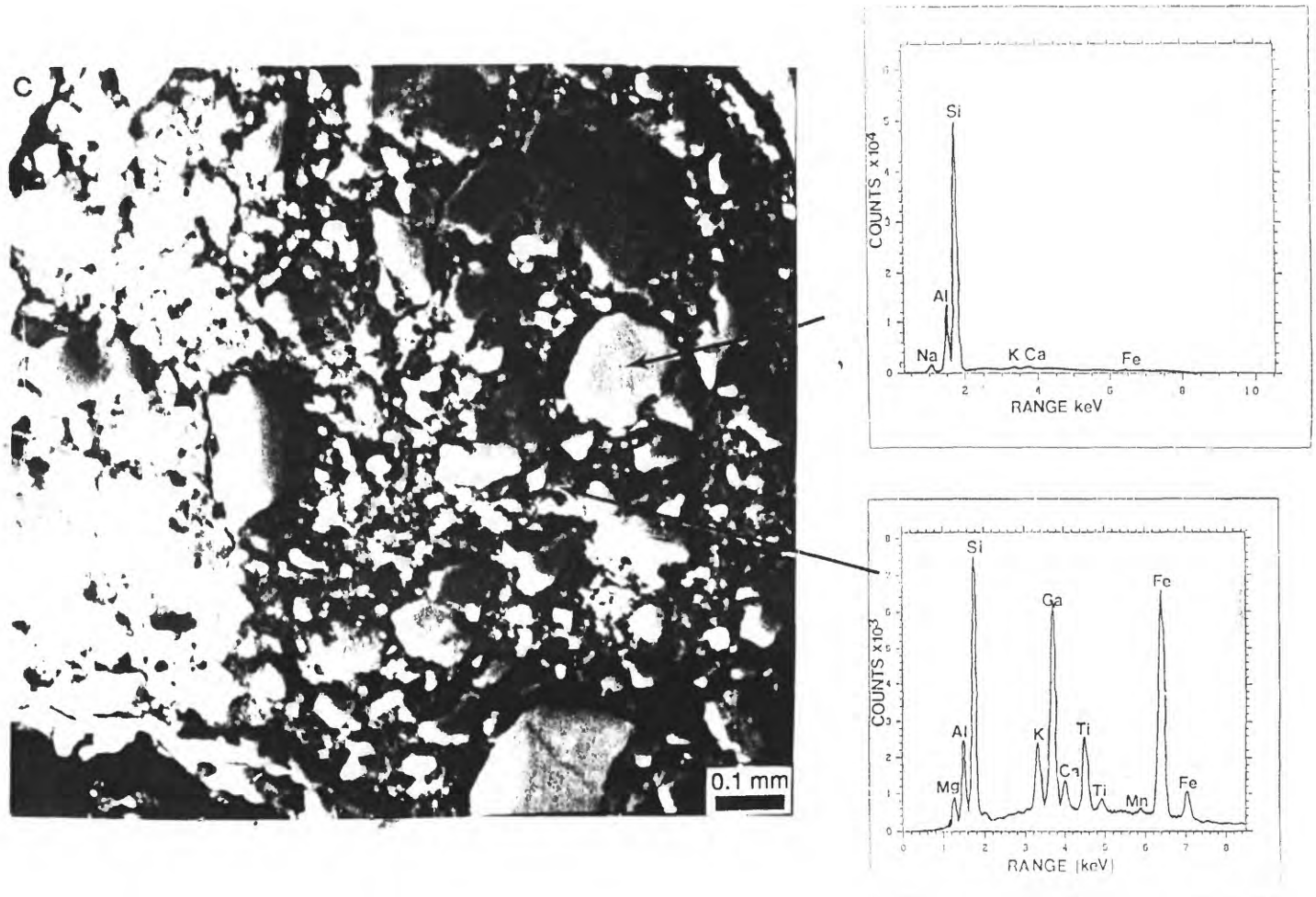



**Figure 3 C.** High resolution back-scattered electron microscope image of fine-grained gouge and subrounded to subangular feldspar fragments. Note lack of foliated fabric in fine-grained gouge. Energy dispersive X-ray analysis of fragments indicate albite composition, whereas fine-grained matrix adjacent to fragments is rich in Mg, Ti, Fe, and Ca. Composition is consistent with Mg-illilte or montmorillonite mixed with Fe-rich zeolite.

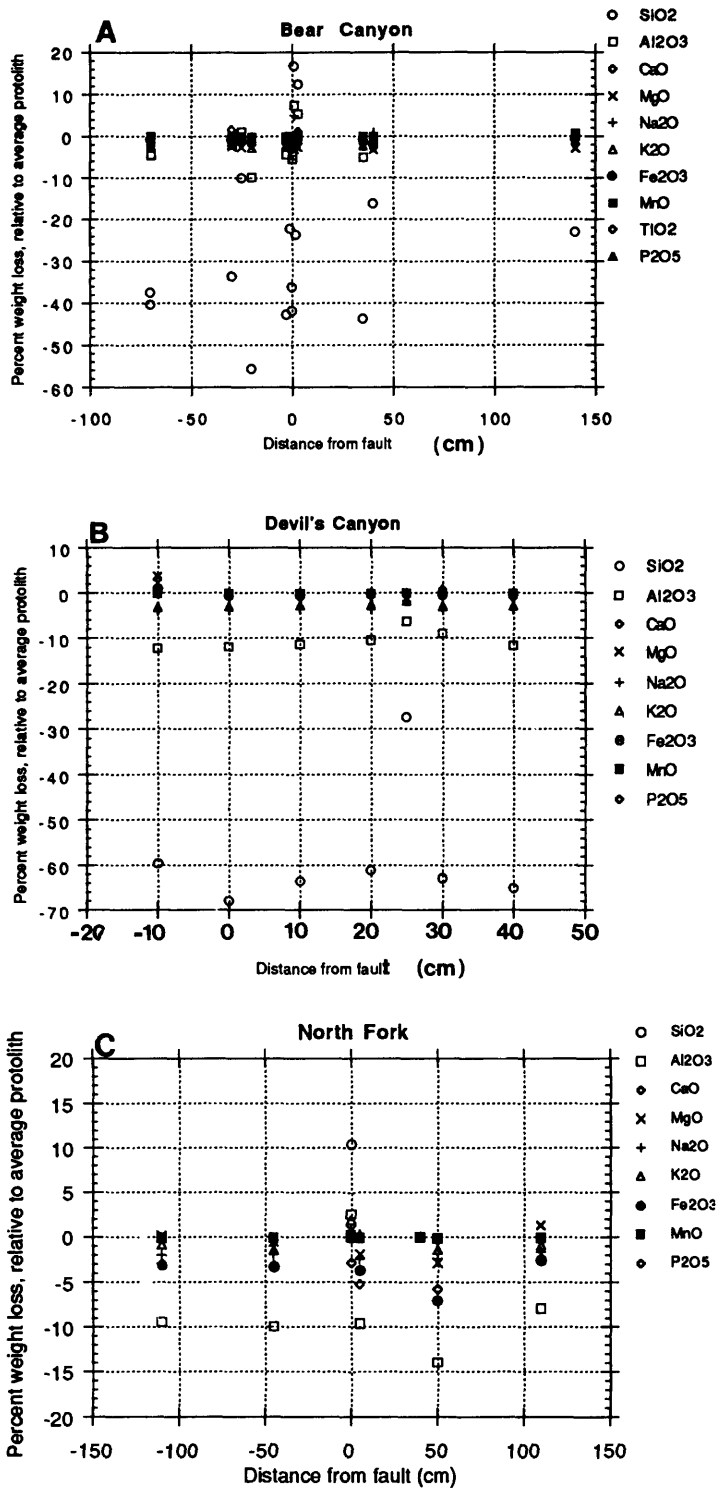



**Figure 4.** Weight loss of each major and minor oxide calculated relative to immobile Ti in averaged protolith as a function of distance across the fault for sample traverses at each site.
A. Bear Canyon.
B. Devils Canyon.
C. North Fork.

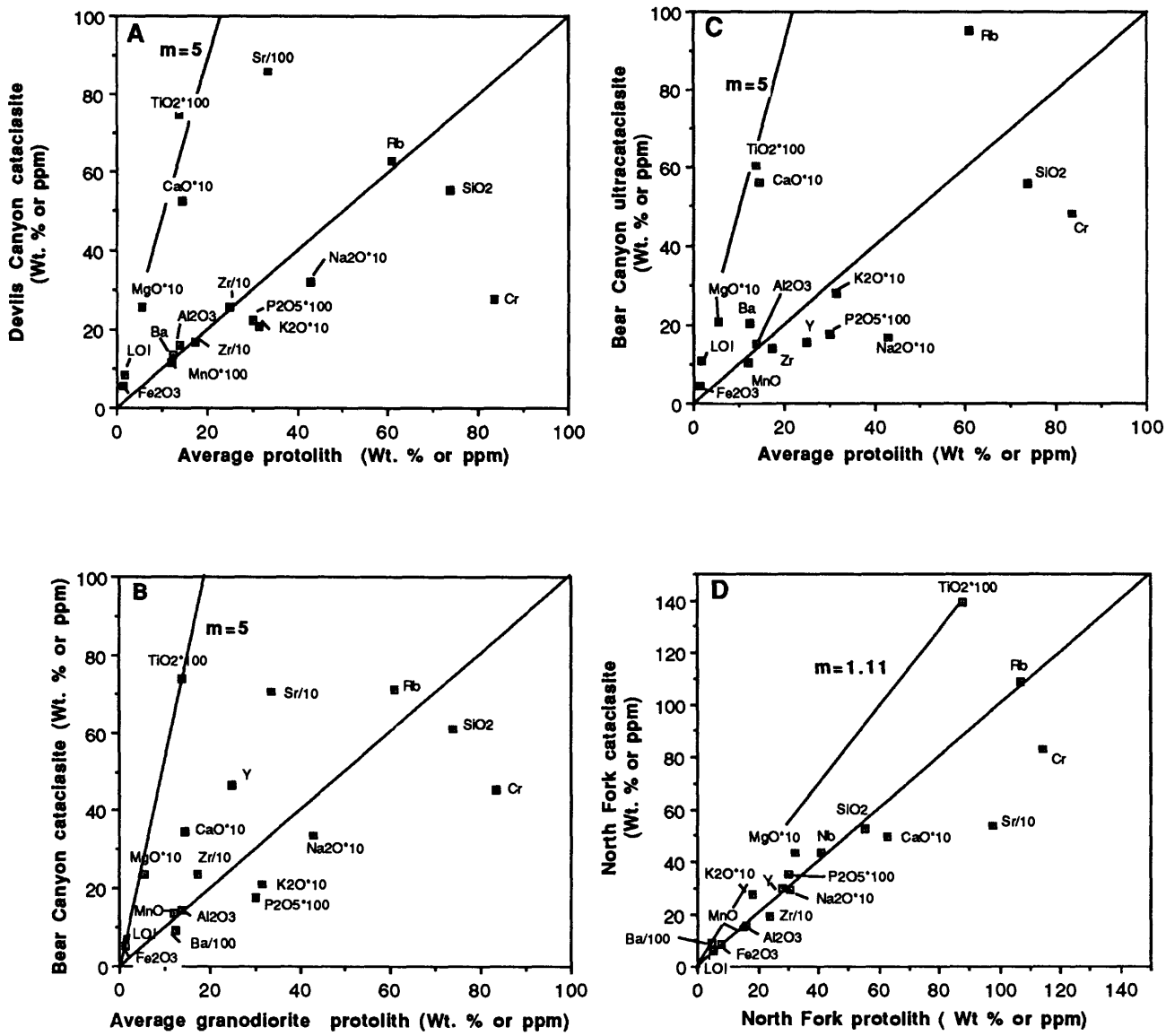



**Figure 5.** Elemental concentrations of averaged cataclasites and ultracataclasites at each site versus average protolith. Line of slope (m) = 1 indicates no loss or gain of elements; elements plotted above the m = 1 are enriched in the deformed sample relative to the protolith; elements which lie below this line are depleted in the deformed sample relative to the protolith (Grant, 1986; O'Hara, 1988). Protolith plotted in A, B, and C is determined from averaging of undeformed samples at Devils Canyon and Bear Canyon, exclusive of amphibolite. Protolith at North Fork site determined from the undeformed diorite at that site. Deformed samples plotted here are averages for cataclasites and ultracataclasites from each site. Volume loss determined from these plots is (1-1/m). Multiplier or divisor of some elemental concentrations enable all data to lie on a common plot, and does not alter the relationships between elements.

A. Cataclasites of North Branch at Devils Canyon site. Values are the average of 6 analyses.
B. Cataclasites of North Branch at Bear Canyon site. Values are the average of 6 analyses.
C. Ultracataclasites of North Branch at Bear Canyon. Analyses are the average of 5 analyses.
D. Cataclasites at North Fork site. Element analyses are the average of 4 separate analyses.

# Water-Rock Reaction In Fault Zones — Reaction Hardening verses Reaction Softening

by Robert P. Wintsch[1], Roy Christoffersen[2] and A.K. Kronenberg[3]

[1]Department of Geological Sciences, Indiana University, Bloomington, IN 47405
[2]SN/Solar System Exploration Div., NASA Johnson Space Center, Houston, TX 77058
[3]Center for Teconophysics, Texas A&M University,College Station, TX 77843-3113

The physical role of high water pressures in lowering fault strength is well known, but the chemical role of water in active fault zones is not well understood. In fact, water is a potentially powerful chemical agent, and reaction between aqueous fluids and minerals in active fault zones are probably universal. This is so because no matter how locally derived, fluids intruding fault zones during a seismic event must change their temperature, pressure, and density, and thus cannot be in chemical equilibrium with the minerals in that fault zone. In long lived fault zones where reactivation of earlier fault splays is common, and where cyclic intrusions of fluids have repeatedly destroyed old minerals and produced new mineral products, these reactions could have an important effect on the mineral mode, and hence the strength of the altered fault rocks remaining in the zone. This contribution explores the variables that affect the compositions of fluids in fault zones, with the goal of predicting the reactions that alter mineralogy, and hence the rheology of modified fault rocks.

## Approach

The role of solid-fluid reactions on fault rock strength can be assessed using aqueous activity diagrams. To model argillaceous rocks and granitoids, the K-feldspar-albite-quartz-$Al_2O_3$-$H_2O$-HCl system is useful. At equilibrium, variables such as $a{K^+}/a{H^+}$ and $a{Na^+}/a{H^+}$ are buffered by three phase assemblages of alkali feldspars, phyllosilicates, and quartz (Wintsch 1975). Consider a fault zone containing a fluid in equilibrium with muscovite, K-feldspar, and albite (pt. A, Fig. 1). If an exotic fluid intruded in the fault zone, it will shift the composition of that fluid away from this equilibrium composition. Dissolution and replacement of some minerals and precipitation of others will move the composition of the fluid back to the original equilibrium composition. The resulting rock would have a different mode, reflecting the operation of this replacement reaction.

## Water and Rock-Dominated Processes

*Water-Dominated Systems.* Consider the intrusion of relatively acidic meteoric water into a shallow fault zone. Its composition plots in the kaolinite field, and the effect of its intrusion and mixing with water already present and in equilibrium with adjacent minerals would be to decrease the pH and shift fluid compositions toward the kaolinite stability field (Fig. 1, arrow to the lower left). If large amounts of water, or water of moderate acidity is intruded, then the composition of the mixed water would reach the kaolinite field (arrow to the lower right, Fig. 1). To reestablish chemical equilibrium, feldspars would dissolve and first kaolinite, and then muscovite would precipitate as the solutions approach chemical equilibrium with the starting assemblage (pt. A, Fig. 1). Reactions would be of the form:

$$3\ \text{K-feldspar} + 2\ H^+(aq) = \text{muscovite} + 6\ \text{quartz} + 2\ K^+(aq) \qquad (1)$$

Water-rock systems such as these in which the composition of the water remains largely unchanged

by reaction are called water dominated. The more fluid present relative to the surface area of reacting rock (i.e. the higher the water/rock ratio), the less the meteoric fluids will be affected by the dissolution of feldspars, and the more phyllosilicates will precipitate. Repeated intrusions of water during the seismic part of the earthquake cycle may lead to the complete replacement of feldspars by muscovite and possibly kaolinite. Because rock strengths may decrease with increasing phyllosilicate content (Shea and Kronenberg, 1993), the resultant phyllosilicate-rich fault zone may be expected to be much weaker than the unreplaced feldspathic protolith before intrusion of the fluid.

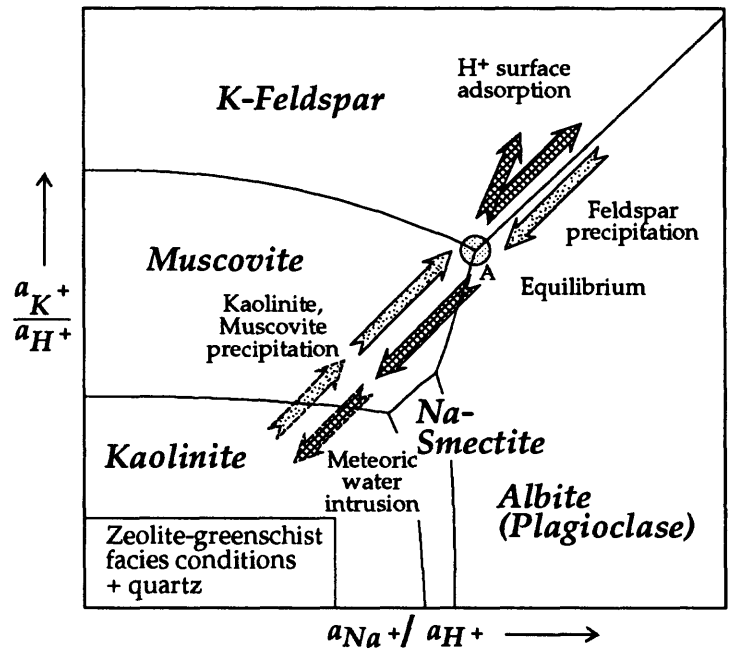



*Figure 1*

*Rock-Dominated Systems.* The converse of the above occurs where the fluid/rock ratio is low. In this environment, the composition of the water is strongly influenced by dissolution, precipitation, and surface exchange reactions in the rock, and the composition of the fluid is said to be rock dominated. The comminution of grains in faults during the seismic event greatly increases the surface area, thus exposing new silicate surfaces to the freshly intruded fluid. Alkali silicate surfaces are strongly $H^+$ selective, and a surface reaction occurs essentially instantaneously:

$$K(\text{feldspar surface}) + H^+(aq) = H(\text{feldspar surface}) + K^+(aq) \qquad (2)$$

In water dominated systems, surface reactions have a very small effect on the pH of the total fluid system. However, if the volume of fluid in the fault is small, then the small amount of $H^+$ in the fluid is rapidly exhausted, and $H^+$ is taken from $H_2O$ to produce alkali hydroxides; potentially very high pH's result. Where water/rock ratios are small, this surface reaction can have a profound effect on variables such as the $K^+/H^+$ and $Na^+/H^+$ activity ratios, and drive the solutions in fault zones strongly away from equilibrium toward more alkaline compositions (Wintsch, 1975; Wintsch and Knipe, 1982; arrow to the upper right, Figure 1). In this case phyllosilicates would dissolve and feldspars would precipitate — the reverse of reaction 1.

**Addition of Magnesium to the System**

The addition of MgO to the above system adds chlorite (modeled as clinochlore) and biotite (modeled as phlogopite) to the potential assemblage (Fig. 2). Albite might still be present in most assemblages, though not explicitly plotted. Muscovite-K-feldspar-bearing assemblages in magnesium-poor rocks are represented by point A in Fig. 2 (equivalent to pt. A, Fig. 1). The intrusion of meteoric water, and the adsorption of $H^+$ on silicate surfaces destroy solid-fluid equilibria in the same way as that described above. In Mg-rich fault zones, however, chlorite joins the assemblage (pt B, Fig 2). In rock dominated systems, a decrease in $a(H^+)$ shifts the fluid into the chlorite field (pt C). At pt. C chlorite precipitation decreases the $Mg^{++}/H^{+2}$ activity ratio, moving the fluid composition

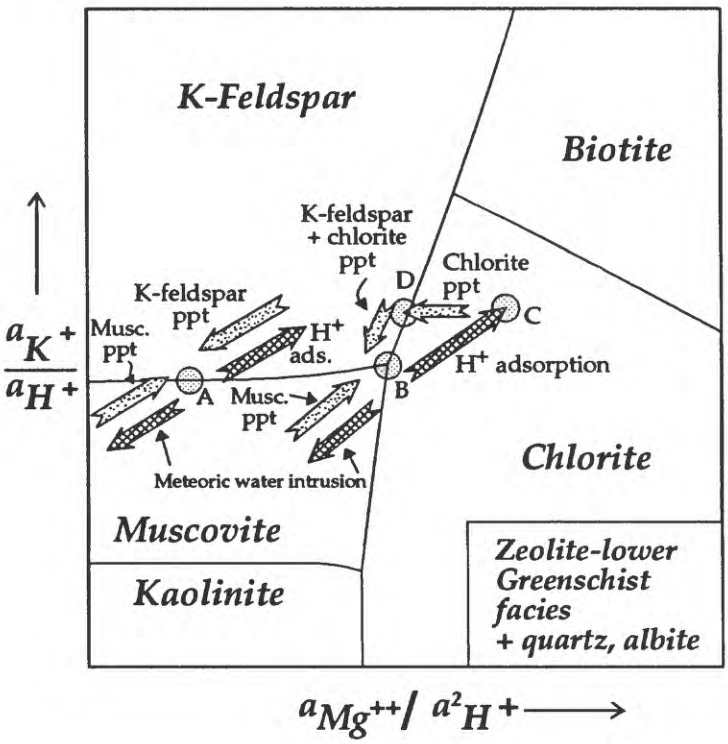



*Figure 2*

to the K-feldspar-chlorite phase boundary (pt. D, Fig. 2). Here K-feldspar + chlorite precipitate simultaneously, until the original equilibrium of pt B is reestablished. These relationships hold for low grade rocks. At metamorphic grades above the biotite isograd, the K-feldspar-chlorite compatibility is replaced by muscovite-biotite equilibria. Biotite replaces chlorite as the magnesian mineral, but the solid-fluid reaction relationships are the same: $H^+$ adsorption still shifts the fluid into the stability field of a phyllosilicate.

**Effect of Reactions on Fault Rock Rheology**

In the water-dominated systems described above, weak phyllosilicates (muscovite and kaolinite) replace stronger feldspars. This process in which strong minerals are replaced by weak phyllosilicates has been called *reaction softening* (White and Knipe, 1978), and faults intruded by meteoric water in multiple earthquake cycles are expected to have reduced shear strength. The converse of this reaction, that can occur in rock dominated granitic systems, destroys weaker phyllosilicates and replaces them with stronger feldspars, thus strengthening the rock. This is an example of reaction hardening (Wintsch and Dunning, 1985), and deeper fault zones are less likely to lose frictional strength.

In fault zones rich in Mg, however, phyllosilicates will precipitate in both water-dominated, and in rock-dominated systems. In shallow, water-dominated fault zones dioctahedral mica (muscovite) will precipitate, whereas in rock-dominated, fluid poor fault zones, chlorite (<10 km) or biotite (>10 km) will precipitate. Consequently, reaction softening should dominate regardless of the permeability of the fault zone. This is in strong contrast to magnesium-poor rocks, where relatively strong feldspars precipitate in deeper, rock-dominated fault systems.

**Depth Distribution of Solid-fluid Reactions in Active Fault Zones**

The concepts outlined above can be applied to active fault zones through a plausible model (adapted from Sibson, 1977) of the distribution of fault rock types as a function of depth and strain rate (Fig. 3). Fast strain rate, random fabric, fault rocks (breccias, cataclasites) may dominate the shallow regions of fault zones, but blastocataclasites are probably produced, though rarely preserved, in prehnite-pumpellyite through amphibolite facies rocks to depths >20 km. Slow strain rate fault rocks include foliated fault gouges at shallower depths, and deeper mylonitic rocks at higher metamorphic grades.

*Water-Dominated Fault Zones.* The intrusion of fluids into fault zones is probably aided by seismic pumping, and would be most strongly favored in rocks deformed during high strain rate, seismic events, where porosity is created (breccias, cataclasites, Fig. 3). In these fault zones

dioctahedral white micas and even kaolinite in the lower zeolite facies may be produced. In more mature fault zones, feldspars may be completely dissolved.

*Rock-Dominated Fault Zones.* At deeper levels consistent with the prehnite-pumpellyite, greenschist and deeper metamorphic facies (Fig. 3), aseismic displacement begins to accommodate an increasing amount of displacement in fault zones, and fault rocks become increasingly well foliated. Here the strain is accommodated by dislocation slip in the micas, especially where micas are aligned in layers (Shea and Kronenberg, 1993), and a mixture of limited dislocation slip and grain-scale microcracking of other phases; very little porosity is created. Thus the intrusion of meteoric water is much more limited, and the compositions of the small volumes of fluids present in deeper, active fault zones may be largely rock-dominated. Because earthquake activity (and grain size reduction) begins to be important at these depths (Sibson, 1982), surface exchange will be important in increasing the pH of the fluids. At these depths, the alkaline fluids produced will precipitate feldspars, and reaction hardening is expected to be increasingly important. The relatively high strength of fault rocks suggested by repeated seismic events in spite of many cycles of brecciation may reflect repeated cementation of fault rocks by relatively strong feldspars.

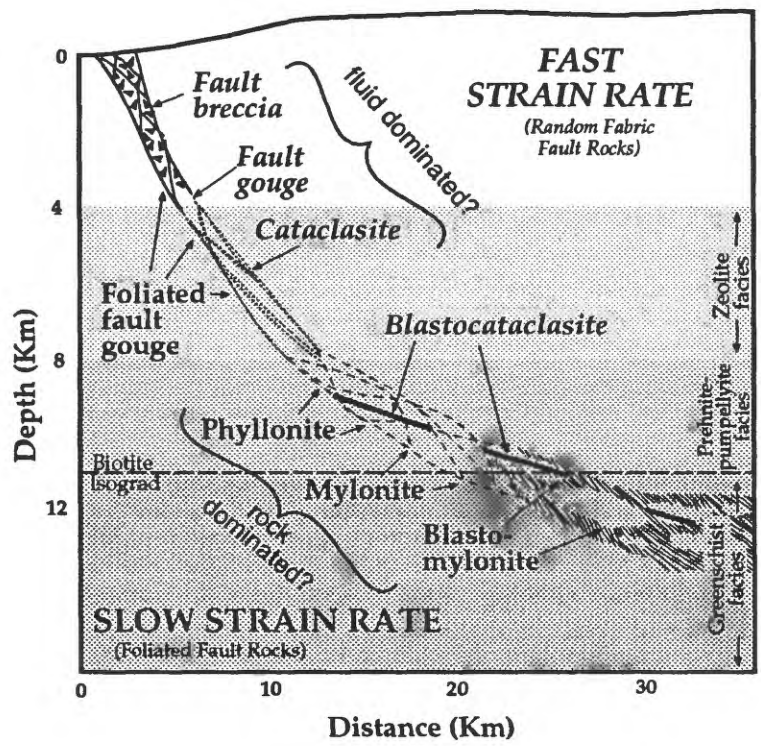



*Figure 3*

*Magnesium-Rich Fault Zones.* An exception to this generalization is found in Mg-rich rocks. Even in prehnite-pumpellyite facies and greenschist facies rocks, chlorite and biotite respectively, may precipitate from alkaline, rock-dominated fluids (Fig. 3). These minerals would keep the fault zone weak, limiting seismic slip, and enhancing aseismic creep. There may even be a feed back, where creep on foliated mylonitic rocks may limit stress levels, reduce the amount of porosity increase by seismic deformation, and hence preserve rock-dominated conditions by retarding the intrusion of meteoric fluids into these deeper faults.

**Shear Strengths of Phyllosilicates**

Several lines of evidence suggest that shear stresses on active crustal fault zones are low compared with the frictional strengths of most crustal rocks (Zoback et al., 1987; Hickman, 1991) and that large-scale fault zones are weak relative to their surroundings over long geologic times. Heat flow measurements across the San Andreas Fault, for example, limit shear stresses to less than 20 MPa, averaged over 14 km of the upper crust (e.g. Lachenbruch and Sass, 1980), while frictional properties of most crystalline rocks (Byerlee, 1978) predict stresses that exceed 100 MPa over much of this same depth interval. Frictional strengths are expected to rise substantially with depth as stresses normal to the fault zone rise unless fluids at near-lithostatic pressures are present. If fault

zones are weak due to high fluid pressures, fluid pressures must remain elevated over extended periods of time. Alternatively, fault zones may be weak due to the presence of phyllosilicates.

Phyllosilicates are weak in comparison with most other silicates over a wide range of conditions and their localized precipitation along fault zones offers a simple explanation of the persistent weakness of mature faults relative to the surrounding crust. Whereas fracture and frictional sliding are expected to govern the mechanical properties of most silicates at shallow to intermediate crustal levels, phyllosilicates deform by dislocation glide at low temperatures and at low shear stresses that are not dependent on mean stress. Basal shear strengths of mica single crystals measured at high mean stresses are small (10 to 30 MPa) at experimental strain rates (Kronenberg et al., 1990; Mares and Kronenberg, 1993); strengths inferred at slow, geologic deformation rates (<10 MPa) are well below those required by heat flow and other geophysical observations.

Experimental studies of mica-bearing rocks indicate that strengths of rocks decrease with increasing mica contents (Shea and Kronenberg, 1993). However, strengths of even the most mica-rich schists (Shea and Kronenberg, 1992) are significantly (3 to 4 times) larger than the basal glide strengths measured for single crystals, since many of the mica grains of an aggregate are poorly oriented for glide. Strengths of mica-bearing rocks within the upper to middle crust, inferred by extrapolating the experimental results for schist to geologic strain rates, exceed the upper bounds established for fault zones. However, direct application of these experimental results may not be appropriate if dissolution and precipitation at grain boundaries over longer, geologic times relax compatibility constraints between neighboring grains and allow deformation of micas that are oriented favorably for glide without requiring much strain in poorly oriented grains. Thus, the strengths of mica-rich rocks may approach those of mica single crystals in weak orientations if dissolution and precipitation assist deformation or if foliations defined by aligned micas are well developed.

**Roughness at the Top of the Seismic Zone**

The reactions predicted above may now help to address the questions of fault rock rheology in active fault zones. In the shallow parts of fault zones, where the intrusion of meteoric water might be routine, fault rocks are predicted to be weakened by the formation of phyllosilicates and shear resistance should be low. Below 8-10 km, however, where the intrusion of meteoric water is more restricted, the compositions of fluids in fault zones are expected to be rock dominated. Here the precipitation of feldspars and replacement of phyllosilicates would strengthen fault rocks. However, the presence of mafic or ultramafic rocks in fault zones may add enough magnesium to the fluids in fault zones that chlorite and biotite would be stabilized. In this case weak phyllosilicates could precipitate from rock dominated fluids to depths well into the greenschist facies, and perhaps even into the lower amphibolite facies where reaction hardening is expected to be more common.

It may be possible that the anomalously low strength of the San Andreas fault zone can be explained, in part, by reaction softening. At shallow, water dominated levels, muscovite may be a common reaction product, while at deeper levels, enough magnesium may be present to allow reaction softening to occur into the greenschist facies. Supporting this idea is the correlation of the occurrence of serpentine (related to ophiolites of the Great Valley sequence) with high creep rates along portions of the San Andreas fault (e.g. Irwin and Barnes, 1975). The implications of this study are that magnesian phyllosilicates may precipitate in many rock types in small coalescing shear zones that make up only a small percentage of the mode of the rock, but if they occur in layers, then their effect on weakening a fault zone may be large. This hypothesis has an important advantage over that

of high fluid pressures. The very low permeability required to sustain high fluid pressures, even through repeated earthquake cycles, is difficult to rationalize. On the other hand, phyllosilicates can lower resistance to shear stress in both high permeability, water-dominated, and low permeability, rock-dominated systems.

## Evidence for the Role of Fluids in the Honey Hill Fault Zone

Several examples of reaction softening and hardening occur in the naturally exhumed Permian-Triassic Honey Hill fault zone in eastern Connecticut. This fault zone contains cataclasites, commonly with K-feldspar-(chlorite)-(sphene) cements, mineralogy that demonstrates the strong alkalinity of the precipitating fluids (Wintsch, 1985). The near monomineralogy reflects water-domination. Multiple brecciation of some cataclasites with increasing chlorite/K-feldspar ratios in succeeding generations of cements is evidence for reaction hardening, and for the growing importance of Mg-bearing fluids in these zones. TEM observations of these K-feldspar cements documents a randomly oriented interlocking network of K-feldspar grains 0.5-2 μm in diameter — a very strong texture (Christoffersen and Wintsch, 1988).

Evidence for reaction softening is also present as slickensided chlorite veins cutting granitic gneisses. The veins are monomineralic, which reflects fluid domination, consistent with the smooth, interpenetrating fracture surfaces in which they occur. The lack of any other minerals, and the local preferred orientation of the chlorite grains meet the criterion outlined above, indicating that this rock had a very low strength, and could have approached that of single crystal phyllosilicates. Monomineralic biotite phyllosilicates are also present in this fault zone. These grains are up to 100 μm thick, 1 mm in diameter, and have a very strong preferred orientation. Again, such layers would have very low strength. If such layers are present in rocks of the San Andreas fault zone, they could account for the weakness of the zone.

## Conclusions

We conclude that phyllosilicate-rich rocks with strong preferred orientations are capable of lowering the strength of mature faults. Phyllosilicates are predicted to precipitate in both rock-dominated and water-dominated systems, provided that magnesium is present. Because magnesium-rich ultramafic rocks are present along parts of the San Andreas fault, the weakness of this fault zone could be explained by the occurrence of phyllosilicate-rich fault rocks. The presence of high fluid pressures may further weaken these rocks, but may not be necessary to explain low strengths.

## References


Byerlee, J.D., 1978, Friction of rocks, *Pure Appl. Geophys.* **116**, 615-626.

Christoffersen, R. and Wintsch, R.P., 1988, The modification of fault rock mineralogy by rock-dominated fluid, Honey Hill fault, Connecticut: *Geol. Soc. America, Abs. with Prog.* **20**, 332.

Hickman, S.H., 1991, Stress in the lithosphere and the strength of active faults, *Reviews of Geophysics* **29**, 759-775.

Irwin, W.P, and Barnes, I., 1975, Effect of geologic structure and metamorphic fluids on seismic behavior of the San Andreas fault system in central and northern California: *Geology* **3**, 713-716.

Kronenberg, A.K., Kirby, S.H., and Pinkston, J., 1990, Basal slip and mechanical anisotropy of biotite, *J. Geophys. Res.* **95**, 19257-19278.

Lachenbruch, A.H., and Sass, J.H., 1980, Heat flow and energetics of the San Andreas Fault zone, *J. Geophys. Res.* **85**, 6185-6222.

Mares, V.M., and Kronenberg, A.K., 1993, Experimental deformation of muscovite, *J. Struct. Geol.* **15**, in

press.
Shea, W.T., and Kronenberg, A.K., 1992, Rheology and deformation mechanisms of an isotropic mica schist, *J. Geophys. Res.* **97**, 15201-15237.
Shea, W.T., and Kronenberg, A.K., 1993, Strength and anisotropy of foliated rocks with varied mica contents, *J. Struct. Geol.* **15**, in press.
Sibson, R.H., 1977, Fault rocks and fault mechanisms, *J. Geol. Soc. Lond.* **133**, 191-213.
Sibson, R.H., 1982, Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States: *Bull. Seism. Soc. Am.* **72**, 151-163.
White, S.H., and Knipe, R.J., 1978, Transformation and reaction enhanced ductility in rocks: *J. Geol. Soc. Lond.* **135**, 513-516.
Wintsch, R.P., 1975, Feldspathization as a result of deformation: *Geol. Soc. America Bull.* **83**, 35-38.
Wintsch, R.P., 1985, Evidence from a fossil fault zone for the strengthening effects of cement in fault gauge (abs.): *Trans. Am. Geoph. Union (EOS)* **66**, 354.
Wintsch, R.P. and Knipe, R.J., 1983, Growth of a zoned plagioclase porphyroblast in a mylonite: *Geology* **11**, p. 360-363.
Wintsch, R.P., and Dunning, J.D., 1985, The effect of dislocation density on the aqueous solubility of quartz and some geologic implications: A theoretical approach. *J. Geophy. Research* **90**, 3649-3657.
Zoback, M.D., Zoback, M.L., Mount, V.S., Suppe, J., Eaton, J.P., Healy, J.H., Oppenheimer, D., Reasenberg, P., Jones, L., Raleigh, C.B., Wong, I.G., Scotti, O., and Wentworth, C., 1987, New evidence on the state of stress on the San Andreas Fault system, *Science* **238**, 1105-1111.

# THE KINETICS OF PRESSURE SOLUTION AT HALITE-SILICA INTERFACES AND INTERGRANULAR CLAY FILMS

Stephen H. Hickman[1] and Brian Evans[2]

[1]*U.S. Geological Survey, 345 Middlefield Road, MS 977, Menlo Park, CA 94205*
[2]*Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139*

## ABSTRACT

To understand better the mechanisms and kinetics of pressure solution, we conducted experiments in which convex halite lenses were pressed against flats of fused silica in brine. Fluid pressures were maintained at 0.1 MPa; temperatures and mean contact normal stresses ranged from 8.3°–90.2° C and 0.5–13.5 MPa, respectively. The geometry and growth rate of the contact spot between the two lenses and the rate at which the lenses approached one another (convergence) were monitored using reflected light interferometry and transmitted light photomicrography. Convergence occurred when halite and silica lenses were pressed together in brine (halite/silica experiments). No undercutting was observed and dry control experiments indicated negligible dislocation creep. Convergence rates in experiments at 50.2° C ranged from 0.01–0.05 μm/day, depending on mean normal stress and contact spot radius. The data were consistent with intergranular pressure solution (IPS) rate-limited by diffusion through a high-diffusivity (c.a. $10^{-5}$–$10^{-7}$ $cm^2/s$) intergranular film. Convergence rates in similar experiments conducted at a constant load of 0.11 N but at 8.3°, 50.2° and 90.2° C were approximately constant at a given normal stress and contact spot size. The cause of this temperature insensitivity is unknown, but it may result from changes in interphase boundary structure or thickness with increasing temperature that are sufficient to offset the expected thermal activation. An experiment was also conducted in which a halite lens was pressed against a fused silica flat coated with an 0.8-μm-thick film of Na-montmorillonite in brine. This clay film produced an approximately five-fold increase in convergence rates over those observed in a halite/silica experiment conducted without clay at the same load and temperature. The strong sensitivity of IPS rates both to contact spot radius and to the presence of second phases along grain boundaries suggests that fine-grained, clay-rich fault gouges and multi-phase granular aggregates should be particularly susceptible to pressure solution creep in the mid- to upper-crust.

## INTRODUCTION

Fluids profoundly affect fault deformation by mechanisms that are both chemical and physical in nature. Stress corrosion cracking, pore-fluid diffusion, and dilatancy hardening on faults may delay sliding instability [*Rudnicki*, 1988]. The influx of deep fluids [*Rice*, 1992], the compaction of porous rocks [*Sleep and Blanpied,* 1992], or the intrinsically low permeability of clays [*Byerlee,* 1990] might cause or permit high pore fluid pressures and reduce frictional strength during seismic events. Compaction of fault rocks could also occur by fracturing [*Wong,* 1990]. Chemical processes such as cementation, sealing, and lithification can indurate and strengthen fault rocks, reduce the permeability of the fault and its surroundings, or increase the elastic stiffness of the fault zone [*Angevine et al.,* 1982; *Chester and Logan, 1986; Fredrich and Evans,* 1992].

Pressure solution may be a particularly important deformation mechanism in the fine-grained rocks and gouges along faults. This process, which involves dissolution of material at grain-to-grain contacts under high stress and the precipitation of that material at interfaces under low stress [*Tada et al.,* 1987], could reduce long-term strength and allow slow, stable fault creep [*Rutter and Mainprice,* 1979; *Chester and Higgs,* 1992]. The fibrous overgrowths and slickolytes observed in many exhumed fault zones are probably evidence of that process [*Durnay and Ramsey,* 1973]. For

example, *Power and Tullis* [1989] described slickensides in fine-grained rocks along a fault in Nevada and inferred cycles of slow, continuous deformation, followed by seismogenic cataclasis. Despite pressure solution's potential importance in fault rheology [*Kirby,* 1980; *Sibson,* 1983], major uncertainties remain regarding the mechanisms and kinetics of this process [*Rutter,* 1983; *Tada et al.,* 1987]. In fact, there is no general agreement even on the form of the appropriate constitutive law. A necessary prerequisite for developing realistic constitutive equations is an understanding of the physics and chemistry of pressure solution at individual grain-to-grain contacts.

Two classes of models have been proposed for intergranular pressure solution: water-film diffusion and undercutting. In water-film diffusion (Figure 1A), material is postulated to dissolve along a loaded grain or interphase boundary under a high normal stress, $\sigma_N$, diffuse through a high-diffusivity film between grains and precipitate in a pore at a pore fluid pressure, $P_f$. Water-film diffusion can be rate-limited either by diffusion in the intergranular film [diffusion limited; e.g. *Weyl,* 1959; *Rutter,* 1976] or by dissolution/precipitation at the solid-liquid interface [interface limited; e.g. *Raj and Chyung,* 1981]. In undercutting models (Figure 1B), the size of the contact spot between grains is thought to be reduced through free-face dissolution of material at the perimeter of the contact, until the strength of the remaining material is exceeded. Failure of the neck region is then presumed to occur, either through brittle failure *[Bathurst,* 1958] or dislocation flow [*Pharr and Ashby,* 1983].

The structure and diffusivity of the grain boundaries (separating two grains of the same phase) or interphase boundaries (separating dissimilar phases) required by these two types of models differ greatly. Water-film diffusion requires a high-diffusivity grain or interphase boundary capable of supporting macroscopic gradients in normal stress. Some workers propose that this high-diffusivity phase is a strongly adsorbed (or structured) water layer, with thermodynamic and transport properties different from those of bulk water [*Weyl,* 1959; *Robin,* 1978; *Rutter,* 1976, 1983; *Tada et al.,* 1987]. Alternatively, others have proposed that this boundary phase consists of load-bearing "islands" surrounded by an interconnected network of fluid-filled channels (i.e. an "island-channel" structure), wherein the fluid has properties close to those of bulk water [*Raj and Chyuang,* 1981; *Spiers and Schutjens,* 1990]. Undercutting models do not require the presence of a high-diffusivity boundary between grains, but they do require either near-zero fluid wetting angles (i.e., the angle formed by the two solid-liquid interfaces where they intersect the grain boundary)[*Pharr and Ashby,* 1983] or enhanced solubility of the material at grain-to-grain contacts [*Bathurst,* 1958; *Weyl,* 1959].

Many experiments designed to investigate pressure solution have involved densification of porous powders or rocks [e.g. *de Boer et al.,* 1977; *Gratier and Guiguet,1986; Spiers and Schutjens,* 1990], but some have studied the strength of low-porosity rocks [*Rutter and Mainprice,* 1978; *Dennis and Atkinson,* 1982; *Urai et al.,* 1986]. In most of these studies, workers have argued that pressure solution occurred based on microstructural observations, and on the stress and grain-size dependence of the densification rate. But, because stress concentrations exist near pores and inclusions, other deformation mechanisms, including dislocation flow and microfracturing, may contribute to strain especially in very porous rocks [see *Dennis and Atkinson,* 1982; *Wong,* 1990]. Microstructural evidence of fracturing can be quickly obscured by dissolution of fines or crack healing. Thus it has been quite difficult to use results obtained from standard high temperature and pressure deformation experiments to judge the merits of theoretical models for pressure solution. In an attempt to remove ambiguities in the identification of deformation mechanisms, a few studies have used simple model geometries [*Sprunt and Nur,* 1977; *Tada and Siever,* 1986; *Gratier,* 1993], although the interpretation of some of these experiments is still controversial [*Bosworth,* 1981; *Green,* 1984; *Gratier,* 1993].

Recently, *Hickman and Evans* [1991, 1992] studied deformation at the contact between halite and fused silica lenses of carefully controlled shape, which were immersed in brine. By observing the model geometry while under load, they were able to distinguish between mechanisms which caused deformation at the contact and those which only caused the growth of the contact zone. When a convex halite lens was pressed against a flat halite lens, no convergence occurred between the two lenses. However, the contact area between the lenses grew as halite dissolved from the free

surfaces of the lenses, diffused through the pore fluid, and precipitated at the perimeter of the contact spot. This process, called neck growth, is analogous to crack healing and is driven by gradients in surface curvature [*Hickman and Evans,* 1992]. In similar experiments with a halite crystal pressed against a flat silica surface, convergence by intergranular pressure solution did occur. Deformation by cataclasis could be ruled out by careful observation of the surfaces and contact region during deformation.

In this paper, we present results from two suites of experiments conducted using halite and fused silica lenses in the same geometry employed in our earlier studies, but in which load and temperature were independently varied. In this manner, we were able to study the kinetics of intergranular pressure solution (IPS) at the halite/silica contacts in greater detail. To investigate the effect that intergranular clays might have upon IPS rates, we also present results from an experiment in which a convex halite lens was pressed against a fused silica flat coated with montmorillonite in brine. The results are a direct experimental confirmation of enhancement of IPS at interphase boundaries, which is often observed in field studies [e.g. *Tada and Siever,* 1986].

## TECHNIQUE

In these experiments, convex halite lenses were pressed against flat silica disks in a heated microscope stage (Figure 2). We review the technique briefly here; additional details are given by *Hickman* [1989] and *Hickman and Evans* [1991, 1992]. To begin each experiment, the halite/silica samples, a spring and fused silica spacers were immersed in saturated brine inside a Teflon chamber, which was surrounded by a constant temperature bath. Special care was taken to insure that the temperature within the sample chamber was spatially uniform and constant in time (within 0.1°C) and that the brine was completely saturated.

Loads were applied using creep- and corrosion-resistant Hastelloy C springs and were deliberately kept small to avoid fracture and minimize dislocation creep. The exact load was determined from the spring compression and a calibration curve. Spring compression was adjusted to the desired value using fused silica spacers of carefully measured dimensions. The calculated load was corrected for thermal expansion of the springs, samples, spacers and sample chamber and for changes in the elastic constants of Hastelloy C. Spring calibrations were checked repeatedly using a dead-weight loading jig equipped with a micrometer touch-point indicator, but no changes in calibration were ever observed. Samples were weighed before and after each experiment and their dimensions were measured using a microscope equipped with a dial indicator.

In experiments at or below 50° C, brine was added directly to the contact spot using a temperature-controlled glass syringe before the chamber lid was closed (i.e. fabricated wet)[see *Hickman and Evans,* 1991]. In the two experiments at 90° C, however, this technique could not be used owing to rapid evaporation of the brine. In those cases, the lid to the sample chamber was closed first, the heated microscope stage was tilted on its side, and the brine was injected into the sample chamber through the lower inlet tube (i.e. fabricated dry; see Figure 2).

### Sample Preparation

The method of sample preparation was important [*Hickman and Evans,* 1991]. Standard abrasive polishing produced thick damage layers on the halite surface that caused roughening when the lenses were exposed to brine. Thus, the samples used here were prepared by cleaving a blank from a single crystal of optical-grade, synthetic NaCl (Harshaw Chemical Company), dissolving this blank to the appropriate size and curvature using porous polymer laps wetted with distilled water, polishing with 0.1 μm alumina grit in isobutyl alcohol using an automatic lens-polishing machine, and baking in air at 650°C for 14 hours. All lenses were polished parallel to [001].

Montmorillonite was used in the halite/clay/silica experiment because of its capacity for adsorbing interlamellar and surface water, even while under high compressive stresses [e.g. *Viani et al.*, 1983]. The Na-montmorillonite film was prepared using standard techniques for preparing basal-oriented clay films for X-ray analysis [e.g. *Gibbs*, 1965]. Montmorillonite from Polksville, Mississippi, (Clay Mineral Standard, API #21, Wards Scientific Co.) was ultrasonically disaggregated and soaked in NaCl brine for about 18 hours to saturate the exchangable cation sites

with sodium. This suspension was repeatedly centrifuged, and the clear brine decanted and replaced with distilled deionized water until flocculation ceased, allowing a fraction with grain size $< 0.5$ μm to be isolated by Stoke's settling. A dilute suspension of this size fraction was centrifuged onto a fused silica disk, and the disk was dried at about 40° C. The sides and bottom of the disk were wiped clean, and the disk was weighed and stored in a dessicator until the experiment. Comparing the weights of the silica disk before and after clay deposition and using a density of 2.5 g/cm$^3$ for montmorillonite, indicated that the clay film was about $0.8 \pm 0.2$ μm thick.

## Measurement of Deformation at the Contact

Convergence was measured using time-lapse photography in reflected sodium light, wherein rays reflected from the disc/brine and brine/lens interfaces produce a circular interference pattern (termed Newton's rings) around the central contact. The distance between the two lenses, $h$, can be determined at any distance,$r$, from the center of contact using

$$h(r) = \frac{m\lambda}{2n_D} \tag{1}$$

where $m$ is the order of an interference minimum, $\lambda$ is the wavelength of sodium light in a vacuum, and $n_D$ is the refractive index of saturated brine for sodium light. We measured $n_D$ at run temperatures using a refractometer connected to a constant-temperature bath [*Hickman*, 1989]. The amount of time-dependent convergence that occurred after loading was determined by monitoring the change in either radius or area of the first ten interference minima surrounding the contact spot using the following relationships [*Hickman and Evans*, 1991]:

$$\Delta s = \frac{\Delta r_m^2}{2R} \tag{2}$$

and:

$$\Delta s = \frac{\Delta A_m}{2\pi R} \tag{3}$$

where $\Delta s$ is the time-dependent (creep) component of convergence; $\Delta r_m$ and $\Delta A_m$ are, respectively, the change in radius and the change in enclosed area of the $m^{th}$ order fringe relative to their values immediately following loading; and $R$ is the radius of curvature of the convex (halite) lens. To minimize distortion, $r_m$ and $A_m$ were measured by projecting the interferogram directly onto a digitizing table. Plots of $m$ against $r_m$ or $A_m$ were used to determine $R$. The mean convergence, $\Delta s$, was then obtained as the mean and standard deviation of values determined independently for each fringe. With the exception of the halite/clay/silica experiment and halite/silica experiment PSH 34, (3) was used to determine convergence in all of these experiments.

## Contact Area and Mean Normal Stress

Contact area, $A_0$, was determined by tracing the innermost indication of reflected sodium light on interferograms projected onto a digitizing table. Uncertainties in contact area were conservatively estimated as the difference between the measured contact area and the area enclosed midway between the contact border and the first interference maximum. The actual precision is considerably better than this. The equivalent contact spot radius, $a$, was defined as the radius of a circle whose area is equal to that of the actual contact. As the contacts in all of our experiments--with the exception of PSH 28 (see below)--were vey close to circular in form, this provides a representation of the contact size that is sufficiently accurate for our analysis. The mean normal

stress, $P_m$, is simply the applied load, $L$, divided by $A_o$. (We subtracted the weight of the lower lens and underlying fused silica disk from the calculated spring force). The spring loads were superposed on the fluid pressure, and so $P_m$ is the effective normal stress (i.e., normal stress minus fluid pressure).

### Modified Tukon Tester

To provide information on the deformation which occurs during initial loading, we modified a Tukon indentation hardness apparatus (Figure 3) to measure $P_m$, as a function of load. A halite lens was prepared using the same procedure as before. The hardness tester was adjusted to lower this lens onto a fused silica flat from a height of about 1 mm in approximately 40 seconds. The contact spot was then photographed under reflected sodium light after a contact time of 2 minutes.

## RESULTS

### Halite/Silica Experiments in the Heated Microscope Stage

Seven experiments were conducted in which halite and silica lenses were pressed together in brine at loads of 0.11–4.22 N and temperatures of 8.3°–90.2° C (Table 1). To ascertain the contribution of dislocation creep to convergence, halite and silica lenses were also pressed together in dry nitrogen at 50.2° C (PSH 18) and 90.5° C (PSH 30) and in moist air at 23°–89.9° C (PSH 34). Except for a power failure during PSH 31 and the room-temperature portion of PSH 34, the maximum temperature excursions were from ± 0.07° to ± 0.20° C. Fortunately, the power failure in PSH 31 had no apparent effect upon the surface topography, contact spot size or convergence rate; nor was there a discernible change in sample weight.

***Contact Morphologies:*** Typical interferograms for a halite/silica experiment in brine at 50.2° C are shown in Figure 4. Notice that the contact spot (central dark region) remains roughly circular, and does not develop the dendritic patterns that epitomize halite/halite experiments [*Hickman and Evans*, 1991, 1992].

Immediate examination of the two experiments fabricated dry at 90° C (PSH 28 and PSH 29) revealed that no air was trapped at the contact in either experiment (Figures 5 and 6). To insure wetting of the contact in experiment PSH 28, the halite lens was then rocked back and forth using a thin wire inserted into the sample chamber. Although this operation succeeded in exposing the contacting surfaces to brine, the halite lens was severely abraded, resulting in an initially "open" (or porous) boundary between the two lenses and a highly irregular contact spot (Figure 6).

Fracturing was not observed during any of the halite/silica experiments (e.g. Figures 4 and 5), with the possible exception of PSH 28 in which the contact region was too damaged to draw conclusions about deformation microstructures.

***Contact Stresses:*** The average contact stress, $P_m$, in the halite/silica experiments immediately after loading ranged from 2.4–12.4 MPa (Figure 7). With the exception of PSH 18 and PSH 30 which were conducted in dry nitrogen, $P_m$ decreases with time as the two lenses converge. The monotonic decrease in $P_m$ with time (i.e., increase in contact spot area) indicates that undercutting did not occur. Note that the initial values of $P_m$ for experiments under comparable loads are quite consistent.

***Convergence:*** To examine the effects of normal stress and contact spot radius, $a$, on convergence rates, halite/silica experiments were conducted in brine at 50.2° C, but at loads of 0.11 N, 0.76 N and 4.22 N. Convergence data are shown in Figure 8; also shown are data from a halite/silica experiment conducted in dry nitrogen at the same temperature and 4.21 N. Note that instantaneous convergence, $s(0)$, occurs immediately after the application of load and is proportional to the applied load. In the experiments in brine, the convergence that follows is indicated by the outward migration of fringes with time (i.e. $A_m$ increases for a given $m$; also see Figures 4-6).

The topography data from the low and intermediate load experiments indicate constant lens curvature (e.g. Figures 8A and B), but the data at high loads indicate a significant deflection of the lens surface near the contact spot ("sinking in"; e.g. Figures 8C and D). Sinking in is indicative of

the plastic deformation of an annealed, work-hardening material [*Matthews*, 1980].

By subtracting the initial elastic/plastic deformation, *s(0)*, from the total convergence, one obtains the time-dependent component of the convergence (henceforth simply called convergence; Figure 9). In PSH 22 and PSH 23, convergence was determined as the mean and standard deviation of Δs using interference minima of orders one to ten. In PSH 17 and PSH 18, because of the sinking in near the contact, convergence was determined using only orders 6–10 and 3–20, respectively. Note that the convergence rates are lowest for the experiment conducted at the highest load (PSH 17) and decrease with time. As explained below, this is a consequence of the competing effects of load and contact spot radius on convergence rates. No convergence was observed in the control experiment conducted at high load in dry nitrogen.

Experiments PSH 22 and PSH 26 were conducted in brine under the same conditions of temperature (50.2° C) and load (0.76 N) to check repeatability (Figure 10). In experiment PSH 26, there was a sharp vertical step of unknown origin in the polished surface of the halite lens. This offset passed so close to the contact spot that it truncated all interference minima above the first order. Thus, the convergence curves for PSH 26 are determined only from the increase in area within the contact spot ($m = 0$) and the first order interference minimum ($m = 1$). In spite of this uncertainty, the convergence rates from these two experiments are similar.

To ascertain the effect of temperature on convergence rates, halite/silica experiments were also done at constant load (0.11 N) but at 8.3°, 50.2° and 90.2° C (Figure 11). A control experiment conducted in dry nitrogen at 0.11 N and 90.5° C indicated negligible dislocation creep. Experiments PSH 23, 29 and 31 show nearly identical convergence rates, even though the temperatures differed by 82° C. Note that the rate from PSH 28, whose contact spot was damaged by the *in-situ* teasing operation, is comparable to that from PSH 29 also conducted at 90° C after the initial period during which scratches and fluid-filled irregularities within the contact disappeared (Figure 6A). Thus, the dry and wet fabrication techniques yield similar boundary transport properties. The initially rapid convergence in PSH 28 might be due to short-circuit diffusion through these fluid-filled irregularities.

Through a variety of processes collectively termed the Rebinder effect, water adsorbed onto the surfaces of ionic and covalent solids can give rise to time-dependent indentation creep [*Latanision*, 1977]. To see if the Rebinder effect contributed to convergence in our experiments, we conducted an additional control experiment in air at high humidity and loads of 4.07–4.23 N. The lenses for PSH 34 were exposed to laboratory air at 41% relative humidity for 27 hours and then sealed in this air in the microscope stage at room temperature. After 188 hours at approximately 23° C, the temperature was increased to 50.2° C for 286 hours and then to 89.9° C for 103 hours, without disturbing the sample. Although an instantaneous convergence accompanied each increase in temperature, no time-dependent convergence was observed (Figure 12).

***Plastic Deformation:*** At the conclusion of PSH 34, the lid to the microscope stage was loosened and tilted, revealing a residual flat spot on the halite lens, with the same size and at the same location as the contact spot (Figures 13A and B). The dislocation distribution was then determined by etching for about 1 minute in a solution of 4-gm $FeCl_3$ per liter of glacial acetic acid, followed by an acetone rinse [*Mendelson*, 1961]. The lens was then split through the contact spot along {001} and etched again. The second etching (Figures 13C and D) revealed very high dislocation densities in a region beneath and surrounding the contact spot. Bands of surface damage elsewhere on the lens surface (Figure 13C) probably resulted from repositioning of the halite lens at the beginning of the experiment.

Since plastic deformation in our experiments extended to depths on the order of several hundred microns (Figure 13D), the lack of a significant Rebinder effect in the present study is not surprising as this effect has been observed only when the plastically deformed zone is confined to depths less than about 3 μm [*Westbrook and Jorgensen*, 1965; *Hanneman and Westbrook*, 1968].

## Halite/Silica Experiment Using the Modified Tukon Tester

In experiment PSH 35, $P_m$ across halite/silica contacts in air was measured as a function of load using the modified Tukon Tester (Figure 14), proceeding from the lowest load (0.098 N) to the highest (10.79 N). The halite lens was mounted on a compliant backing and tilted with

increasing load. The resulting shift in position of the contact spot caused flattening of the halite lens that was first evident at 0.98 N but became more severe as load increased. By 1.96 N, the diameter of the plastically flattened spot was roughly equal to the entire contact spot. Also shown in Figure 14 is the calculated $P_m$ for the purely elastic contact between a sphere of radius 8 cm and a flat plate, assuming no friction [*Timoshenko and Goodier*, 1970], using room-temperature elastic properties for halite (average of Voigt and Reuss averages [*Simmons and Wang*, 1971]) and fused silica (Esco Products Inc., 1986). Note that the agreement between $P_m$ measured at the beginning of the experiments conducted in the microscope stage and the data from the modified Tukon tester is good, especially at low to intermediate loads. The initial $P_m$ from PSH 17, 18 and 34 are somewhat lower than the Tukon data at similar loads. This is probably due to a combination of plastic flattening at high loads and the temperature dependence of $P_m$.

## Clay Film Experiment in the Heated Microscope Stage

***Contact Morphologies:*** A halite/clay/silica experiment (PSH 27) was conducted in brine at 50.2° C and 0.76 N (Table 1). During this experiment $P_m$ decreased from about 2.79 MPa to 0.59 MPa in 359 hours (Figure 7B). The sample was fabricated dry, and some small air bubbles were trapped within the central contact spot at the beginning of the experiment. These completely disappeared within 2 hours after the addition of brine. The contact spot grew rapidly and retained a roughly circular geometry (Figure 15). The dark patch above the contact spot in Figure 15A resulted from migration of the contact spot during the first 15 hours of the experiment in response to a gradual rocking of the halite lens, most of this migration occurring in the first 2 hours.

Fine fibrous extensions of the central contact spot disrupted low-order fringes, making it impossible to determine the order of interference minima simply by counting outwards from the central contact spot. Therefore, a variable wavelength monochromatic light source was used to identify order numbers during this experiment. By measuring the change in wavelength required to offset an interference minimum at a given location by exactly one order number, we obtained two combinations of $\lambda$ and order number, i.e. $\lambda_1 m$ and $\lambda_2(m + 1)$; to eliminate $h$ from Equation 1 and determine $m$ directly.

Use of the variable wavelength monochromatic illuminator after 359 hours (Figure 15B) showed that the lowest order minimum visible was $m = 1$, corresponding to a lens separation of 0.214 μm. Given the approximately 0.8 μm thickness of the clay layer, the presence of interference minima of $m < 4$ adjacent to the contact spot indicates that these minima originate from reflections off the halite/brine and brine/montmorillonite interfaces, and not from the halite/brine and montmorillonite/silica interfaces. This conclusion is supported by calculations of the relative intensities of reflected light using Fresnel's formula [*Rossi*, 1957], which show that reflections from the halite/brine and brine/montmorillonite interfaces should be about one order of magnitude brighter than reflections from either the halite/montmorillonite or montmorillonite/silica interfaces. These low reflectances also explain the absence of Newton's rings *within* the central contact spot (Figure 15).

***Convergence:*** Equation 2 was used to determine convergence in the halite/clay/silica experiment using orders 5-14 (Figure 16). To avoid errors in convergence due to the contact spot migration, the radius of each fringe was determined from the maximum diameter that could be drawn parallel to a ray constructed at 1.57 hours. Notice that $s(0)$ from this experiment, once multiplied by $\pi$ to convert from $r_m^2$ to $A_m$, agrees quite well with $s(0)$ from the halite/silica experiment conducted at the same load and temperature (Figure 8B).

The convergence from the halite/clay/silica experiment is shown in Figure 17, after subtracting the initial convergence, $s(0)$. Also shown are convergence curves from halite/silica and halite/halite experiments conducted in brine at comparable temperatures and loads [*Hickman and Evans*, 1991]. Comparing the halite/clay/silica experiment (PSH 27) with the halite/silica experiment (PSH 22) indicates a nearly five-fold increase in convergence rates due to the presence of an approximately 0.8-μm-thick montmorillonite layer.

## DISCUSSION

### Mechanics of Initial Elastic-Plastic Contact

Even at the lowest loads (≈0.1 N), $P_m$ is less than that calculated for purely elastic contact (Figure 14) [*Johnson*, 1987], suggesting that the initial deformation of the contact region included some plastic yielding in all the experiments. When a rigid sphere indents an elastic-plastic half space, plastic slip initiates at a depth beneath the contact about equal to the contact spot radius at $P_m \approx 1.1Y_s$ [*Johnson*, 1987], where $Y_s$ is the single-crystal yield stress for loading along (001). For the halite used in these experiments, $Y_s$ ranges from 0.6–3.0 MPa [*Davidge and Pratt*, 1964; *Davis and Gordon*, 1968; *Carter and Heard*, 1970]. With increasing load, the plastic zone under the indentor expands until it breaks out to the free surface. In an annealed, work hardening material (such as the halite used here), $P_m$ will then increase slowly with increasing load after this fully plastic mode of deformation is reached.

At a load of about 2 N, the plastically flattened spot in PSH 35 was observed to be about the same size as the contact spot. This plastic flattening, the slow increase in $P_m$ with load above about 3 N (Figure 14), and the high dislocation densities observed in PSH 34 at 4.2 N (Figure 13) indicate that fully plastic deformation was achieved at loads greater than 2–4 N.

The distribution of normal stress acting on the grain/grain interface is important for modeling IPS. When an elastic sphere and flat are in contact, the effective normal stress, $\sigma_N - P_f$, is greatest at the center of the contact (where $\sigma_N - P_f = 1.5\ P_m$) and decreases to zero at the perimeter. This Hertzian stress distribution is truncated when plastic yielding occurs [*Johnson*, 1987]. Deformation during the initial convergence in the halite/silica experiments in the heated microscope stage covers the full range of elastic-plastic to fully plastic behavior, thus the initial stress distributions across the contacts in our experiments are unknown. Further, the initial stresses will evolve towards a new steady-state distribution as elastic, plastic, and diffusive mass-transfer processes occur.

### Kinetics of Pressure Solution: Normal Stress and Path Length Dependence

Pressure solution is driven by gradients in the chemical potential induced by non-hydrostatic stresses. The thermodynamic formulation of this problem has often been debated [e.g. *de Boer*, 1977; *Robin*, 1978; *Lehner*, 1990]. *Paterson* [1973] and others have used the concept of local equilibrium within a grain or interphase boundary to derive an expression for the change in chemical potential owing to deviatoric stresses:

$$\Delta\mu = (\sigma_N - P_f)V \tag{4}$$

where $\Delta\mu$ is the work done in transferring one mole of atoms from a grain or interphase boundary under a normal stress $\sigma_N$ to the wall of the open pore under a fluid pressure $P_f$ and $V$ is the molar volume of the solid. Chemical potential gradients arising from changes in internal strain energy or molar volume are assumed to be small [*de Boer*, 1977; *Tada et al.*, 1987].

Intergranular pressure solution via water-film diffusion requires the operation of three sequential processes: dissolution into a grain or interphase boundary film under high normal stress, diffusion through that film, and precipitation at interfaces under low normal stress. The slowest of these three steps will determine the overall deformation rate. Below, we compare the results from our halite/silica experiments in brine with two competing models for the kinetics of intergranular pressure solution. These models, which were derived for the diffusion creep of polycrystalline aggregates, have been modified for the special case of IPS across a single grain-to-grain contact. In accord with our experimental geometry, these contacts are assumed to be plane and circular, under a mean normal stress, $\sigma_N$, and surrounded on all sides by pores containing fluid at a uniform pressure, $P_f$ (see Figure 1).

Models for IPS rate-limited by diffusion within a thin intergranular fluid film have been presented by a number of workers [e.g. *Weyl*, 1959; *Rutter*, 1976; *Spiers and Schutjens*, 1990]. Two assumptions are commonly made: 1) there is no internal deformation of the grains during

pressure solution and 2) the thickness, $\delta$, and diffusion coefficient, $D_b$, of the intergranular film are independent of normal stress. For our geometry, these assumptions require that the dissolution rate into the intergranular film is constant throughout the contact spot and that the normal stress distribution across the contact spot is parabolic. Then, using Equation 4, the convergence rate, *ds/dt*, is given by [*Rutter*, 1976]:

$$\frac{ds}{dt} = \frac{8C_o D_b \delta V^2 P_m}{R T a^2} \tag{5}$$

where $C_o$ is the solubility (moles solute per unit volume solution) of the unstressed solid in the fluid, $P_m$ is the mean effective normal stress, $R$ is the gas constant, $T$ is the absolute temperature and $a$ is the radius of the contact spot. This model predicts IPS rates across an individual contact that vary linearly with mean effective normal stress and inversely proportional to the square of the contact spot radius. Note, however, that the linear stress dependance implied by (5) holds true only if there is no coupling between $D_b\delta$ and $\sigma_N$ (i.e. the boundary is rigid and incompressible). This assumption has been challenged by a number of authors [*Weyl*, 1959, *Rutter,* 1983; *Tada et al.*, 1987] who have postulated decreases in both $\delta$ and $D_b$ with increasing normal stress. If true, creep rate would be less sensitive to stress than (5) suggests.

Alternatively, *Raj and Chyung* [1981] and *Raj* [1982] have suggested that the diffusivity of an intergranular fluid film may be so high that the rate of IPS would be controlled by the rate of transfer of the solid across the solid/fluid interface. To maintain the high-diffusivity intergranular boundaries required in their model, they propose a thermodynamically stable island-channel boundary structure and near-zero fluid wetting angles. Assuming that either dissolution or precipitation kinetics are rate-limiting (i.e., control the overall rate of IPS), the convergence rate across an individual grain-to-grain contact is [*Raj*, 1982]:

$$\frac{ds}{dt} = \frac{K_+ \chi V P_m}{R T} \tag{6}$$

where $K_+$ is a kinetic constant for either dissolution or precipitation (whichever is slowest) and $\chi$ is the solubility (mole fraction) of the solid in the solution. For deformation across a single contact, this model predicts convergence rates that are linear in effective normal stress and independent of the contact spot radius.

To test the dependence of convergence on path length and normal stress, we determined the time-averaged convergence rate, $\Delta s/\Delta t$, and the time-averaged, mean effective normal stress, $\overline{P_m}$, from the halite/silica experiments in brine at 50.2° C and at loads of 0.11 to 4.22 N. During these experiments $\overline{P_m}$ ranged from 0.53–11.9 MPa and the halite and silica lenses converged at 0.01–0.05 μm/day (Figure 18). As the contact area increased, $\overline{P_m}$ and $\Delta s/\Delta t$ decreased monotonically (Figure 7). Contrary to the predictions of the interface model (Equation 6) the convergence rates did not depend solely upon $\overline{P_m}$. Rather, there are sharp differences in convergence rates between different experiments at similar normal stresses. Compare, for example, the end of PSH 22 with the beginning of PSH 23, where $\Delta s/\Delta t$ *increases* by about a factor of 4 as $\overline{P_m}$ *decreases* from 2.3 to 1.8 MPa (Figure 18). Importantly, this discontinuity in $\Delta s/\Delta t$ coincides with a decrease in the mean contact radius from 325 to 152 μm. This inverse relationship between convergence rate and contact spot size clearly indicates that the interface-limited model is not applicable.

To test the diffusion-limited model, we adjusted convergence rate to be $(\Delta s/\Delta t)\,\bar{a}^2$ and plotted this against $\overline{P_m}$ (Figure 19). The adjusted rates now overlap and are roughly linear in $\overline{P_m}$. Thus, Equation 5 appears to predict the stress and path-length dependence correctly and the data are consistent with a model for IPS rate-limited by fluid film diffusion. Note, however, that a straight line fit to the data does not pass through the origin, suggesting that $D_b$, $\delta$, or both may be increasing rapidly when $\overline{P_m}$<1 MPa. The linear dependence of the corrected convergence rate on mean normal stress agrees with recent experiments by *Gratier* [1993], which suggested a linear relationship between indentation rate and $P_m$ for halite single crystals indented by a cylindrical

stainless steel indentor in brine.

*Rutter's* [1976] model can be used to estimate the interphase boundary diffusivities. Using (5), our data indicate that the product $D_b\delta$ ranges from 1.3 X $10^{-12}$ $cm^3/s$ to 1.8 X $10^{-13}$ $cm^3/s$ when $\bar{P}_m$= 0.56–10.8 MPa. Although it is impossible to solve independently for $D_b$ or $\delta$, estimates for $\delta$ can be used to bound $D_b$. A lower bound on $\delta$ is 1 nm, about 4 monolayers of water. A rough upper bound was obtained by interpolating between adjacent low-order interference minima to estimate the thickness of the brine layer immediately outside of the central contact spots (e.g. at the innermost trace of reflected sodium light in Figures 4–6). The estimates ranged from 19–29 nm, so we infer the intergranular film thickness to be less than 30 nm. A similar upper bound is obtained by assuming that this intergranular film has a refractive index similar to that of bulk brine and noting that the central contact spots were uniformly dark under reflected sodium illumination. Ignoring for the moment any possible variations in $\delta$ or $D_b$ with normal stress, this indicates that $D_b$ is between 1.3 X $10^{-5}$ $cm^2/s$ and 6.1 X $10^{-8}$ $cm^2/s$. These values are many orders of magnitude greater than diffusion coefficients measured along solid state grain boundaries at much higher temperatures and are about the same as diffusion coefficients of electrolytes in brine at 50° C (about 2.6 X $10^{-5}$ $cm^2/s$ [*Vitagliano*, 1960]).

## Kinetics of Pressure Solution: Temperature Dependence

To test the dependence of convergence rate on temperature, we also computed $\Delta s/\Delta t$ and $\bar{P}_m$ from the halite/silica experiments conducted at constant load (0.11 N) and at temperatures between 8.3°–90.2° C (Figure 20A). We did not use PSH 28 in this analysis because of its highly irregular contact geometry. As with the constant-temperature experiments (Figure 18), both the convergence rate and the mean normal stress decreased with time during the constant-load experiments. Convergence rate dropped from 0.05 μm/day at 1.7 MPa to 0.01 μm/day at 0.5 MPa. Except for slightly greater rates in the 8.3° C experiment, the rates are roughly invariant with temperature at a given normal stress. Even after correcting these data for path-length effects as required by the diffusion-limited model (Equation 5), there is no systematic variation in convergence rates with temperature (Figure 20B).

The lack of significant temperature dependance is quite surprising in light of the thermally activated nature of the diffusion process in both liquids [e.g. *Bockris and Reddy, 1970*] and solids [e.g. *Shewmon*, 1983]. Although it is often thought that diffusion within structured intergranular fluid films should be slower than in bulk liquids [*Rutter*, 1976; *Tada et al.*, 1987], there is virtually no theoretical or experimental basis upon which to estimate an activation energy for thin-film diffusion. Nevertheless, following the suggestion by *Rutter* [1976] that the temperature sensitivity of water-film diffusion might be related to that for the viscous flow of water (via the Stokes-Einstein equation) and the heat of solution of the diffusing species, it is of interest to compare our results with diffusion kinetics in bulk electrolytes.

The coefficient appropriate for coupled chemical diffusion of $Na^+$ and $Cl^-$ in brine is the mutual diffusion coefficient, $D^v$. Measurements of $D^v$ have been made for NaCl in brine at temperatures of 0°–50° C and concentrations up to 5.4 M [*Vitagliano*, 1960]. Extrapolating these data to saturation and assuming an Arrhenius dependence of $D^v$ on temperature yields an activation energy of 18 kJ/mol. If the activation energy for interphase boundary diffusion is comparable to that for bulk brine, if $\delta$ does not depend upon temperature, and incorporating the small variation in $C_o$ with temperature [*Int. Crit. Tables, 1928, v. III*, p. 105] then $(\Delta s/\Delta t)\,\bar{a}^{\,2}$ should have increased by a factor of 4.5 as temperature increased from 8° to 90° C. This simple view of the transport properties of the interphase boundary film in our experiments is clearly incorrect (Figure 20B).

Although we believe we have correctly identified the rate-limiting step for IPS in these experiements, it is worth noting that the interface-limited model (Equation 6) should produce an even greater temperature sensitivity than that calculated above. We are aware of no direct measurements of dissolution or precipitation rate constants for halite in water, but *Olander et al.* [1982] estimated an activation energy for halite dissolution of about 50 kJ/mol using data from observations of the migration of small brine inclusions in halite in a temperature gradient. Using this activation energy in Equation 6 suggests that the convergence rate in our experiments would have increased by about two orders of magnitude from 8°–90° C if the kinetics were interface-

controlled.

Rock textures diagnostic of pressure solution are observed over a wide range of pressure/temperature conditions in low grade metamorphic environments [e.g. *Trurnit*, 1968; *Elliot*, 1973; *Engelder and Marshak*, 1985; *Tada and Siever*, 1989], suggesting that the activation energy for this process is small [see *Rutter*, 1983]. The difficulty of isolating IPS in laboratory experiments, however, has meant that experimental determinations of an activation energy for this process have been somewhat equivocal. Still, some estimates of activation energy for IPS are available for comparison with our results. *Spiers et al.* [1990] studied uniaxial compaction creep of synthetic halite aggregates in brine at 20°–90° C, uniaxial stresses of 0.5–2.2 MPa and atmospheric fluid pressure in a piston-die apparatus. In those experiments, sieved reagent grade NaCl powder (grain size = 100–275 μm) was compacted dry at 2.1 MPa for about 15 minutes, producing starting porosities of about 42%. Samples were then saturated with brine and densified in experiments lasting about ten days. Volumetric strain rate was approximately linearly proportional to effective applied stress and inversely proportional to grain size cubed [*Spiers and Schutjens*, 1990]. Comparing the mechanical data with theoretical predictions, those authors concluded that densification occurred via diffusion-controlled IPS, with an activation energy of 24 kJ/mol. Although *Spiers et al.'s* data show a clear increase in compaction rate with increasing temperature, it must be emphasized that their data are from monomineralic halite aggregates.

*Pharr and Ashby* [1983] reported an activation energy of 19 kJ/mol for water-enhanced creep of potassium chloride. However, given the high initial porosities in their samples and the accelerating tertiary creep rates observed in their uniaxial creep tests, it is difficult to rule out cataclasis as contributing to the enhanced creep rates. Even if only IPS occurred in their experiments, no microstructural observations were made, nor was the grain size dependence of creep rate determined; thus, one cannot use their data to differentiate between the diffusion-limited, interface-limited and undercutting models. Because both *Pharr and Ashby* and *Spiers et al.* investigated monominerallic aggregates, the relevance of the activation energy determined in their experiments to water-film diffusion at interphase boundaries--such as those studied in our halite/silica experiments--is unclear.

## Interphase Boundary Structure and Properties in the Halite/Silica Experiments

Even though IPS clearly occurred in our halite/silica experiments, there is no evidence, at least on an optical scale, for metastable island-channel structures as postulated by *Raj* [1982], *Spiers and Schutjens* [1990] and *Lehner* [1990]. Those workers theorized that island-channel structures should form in response to highly localized gradients in the Helmholtz free energy (including contributions from elastic and plastic deformation) along grain/interphase boundaries. Although *Spiers et al.* [1990] interpreted microstructural observations made after their halite compaction experiments as indicating the presence of a fine-scale island-channel network along the grain boundaries during deformation, they note that structures could be modified during or after unloading. The absence of these structures in our experiments is especially significant because we were able to observe the halite/silica interphase boundaries directly while under load, and, as shown above, the contact regions in our high-load halite/silica experiments initially deformed in a fully plastic manner, with very high dislocation densities in direct contact with the interphase boundary.

As there was no undercutting in our halite/silica experiments, they provide a direct demonstration of the water-film diffusion mechanism for IPS. In particular, these experiments indicate that the boundaries between dissimilar materials in aqueous fluids can possess remarkable transport properties, and that these properties, in turn, may exert a profound influence on the kinetics of IPS. Below, we briefly discuss the nature and properties of thin fluid films along interphase boundaries.

***Forces Between Surfaces:*** The forces acting between surfaces contribute to such diverse processes as subcritical crack growth [*Wiederhorn and Fuller*, 1989] and the stability of colloids [*Hiemenz*, 1986]. In addition to the strong chemical bonds that act at very small distances (< atomic dimensions), the dominant long-range (< several tens of nm) forces acting between two surfaces in an aqueous solution are van der Waals forces; hydration, or solvation, forces; and

forces arising from overlap of the diffuse portion of the electrical double layers asociated with the adjacent surfaces [*Bockris and Reddy*, 1970; *Adamson*, 1982; *Israelachvili*, 1985a]. The sum of these forces, called the disjoining pressure, will determine the thickness of a fluid film between two solids in equilibrium with an applied normal stress. As our experiments were conducted in a concentrated electrolyte solution, the diffuse portion of the electrical double layer (the Gouy-Chapman layer) is expected to be extremely thin (The Debye length is about 0.1 nm in saturated brine at 50° C.). Thus, disjoining forces due to double layers can probably be neglected.

In general, the largest of the van der Waals forces are dispersion forces, which arise through electromagnetic interactions between fluctuating dipoles in adjacent solids. Dispersion forces between two plane-parallel surfaces decay as the inverse third power of plate separation. For two bodies of the same material separated by a liquid, the total van der Waals interaction is always attractive. Van der Waals interactions between dissimilar materials in a liquid, however, may be either attractive or repulsive depending upon the relative dielectric properties of the two solids and the intervening liquid. From the approximate solution derived by *Israelachvili* [1985a, Equation 11.13], we calculate that the van der Waals interaction should be attractive in all of our experiments in brine, but that it should be twice as strong between two halite lenses as between a halite and a silica lens.

Even when there is a strong van der Waals attraction between two surfaces, adhesive contact can be prevented by hydration forces. Although hydration forces are poorly understood, they are thought to result from the structuring and ultimate expulsion of hydrated ions and water molecules as two hydrophilic (water-loving) surfaces are brought together [*Etzler and Drost-Hansen*, 1983, and *Israelachvilli*, 1985b]. Hydration forces have been measured between curved surfaces of mica [*Pashley*, 1981, 1985; *Pashley and Israelachvilli*, 1984] and fused silica [*Peschel and Adlfinger*, 1969, 1971; *Peschel et al*, 1982] in pure water and in electrolyte solutions. These measurements show that hydration forces 1) are always repulsive; 2) dominate over other surface forces at separations, $\delta$, less than about 5–10 nm; 3) increase exponentially with decreasing surface separation, with a superimposed oscillatory component for $\delta < 1.5$ nm; and 4) have characteristic decay lengths in electrolyte solutions that are relatively constant at 0.6–1.1 nm for concentrations greater than about 0.01 M. Unfortunately, we are aware of no measurements of disjoining pressure for either the halite/silica or halite/halite systems in saturated brine.

***Conditions for the Existence of Thin Fluid Films:*** Consideration of these surface forces in conjunction with strong (but short range) chemical bonds yields a simple qualitative explanation for the difference in behavior between the halite/halite experiments in brine, for which no convergence was observed [*Hickman and Evans*, 1991] and the halite/silica experiments in brine, for which convergence was observed (this study). Two processes must occur in sequence before two particles can be brought together to form a solid-state grain or interphase boundary [*Clarke*, 1987a]. The first, homogeneous thinning of the intervening fluid layer, is governed by the net disjoining pressure between the two phases. Although the relative magnitudes of the hydration forces appropriate for these experiments are unknown, the van der Walls interaction alone would tend to promote a thinner intergranular fluid layer in the halite/halite experiments. This prediction is consistent with observations and theory indicating that intergranular glass phases in ceramics are typically thicker between dissimilar materials than they are between similar materials [*Clarke*, 1987a].

To remove a film of thickness $\delta$ the adsorbed fluid layers must be expelled to form a grain boundary. This process of grain boundary formation (or dewetting) is driven by the change in overall interfacial energy due to the replacement of two solid/liquid interfaces with an island (of height $\delta$) containing a single solid/solid interface. If $\delta$ is greater than zero, dewetting will require the nucleation and growth of coherent or semi-coherent grain boundaries at isolated contact points. The energy barrier, $\Delta E^*$, for nucleation of an island of dewetted boundary is approximately [*Clarke*, 1987b]:

$$\Delta E^* = \frac{\pi \gamma_{sl}^2 \delta^2}{2\gamma_{sl} - \gamma_{ss}} \quad (7)$$

where the solid/solid and solid/liquid interfacial energies, $\gamma_{ss}$ and $\gamma_{sl}$, respectively, are assumed to be isotropic.

Two conclusions are apparent from (7): dewetting will not occur if the grain-boundary wetting angle, $\theta = 2arccos(\gamma_{ss}/2\gamma_{sl})$, is zero; and the nucleation barrier will increase with increasing fluid film thickness. Observations of fluid inclusions and boundary structures along halite grain boundaries in brine indicate that $\theta$ in the halite/halite experiments was non-zero (i.e. $\gamma_{ss}/2\gamma_{sl} \approx 0.9$) [*Hickman and Evans*, 1992]. In those experiments, the nucleation step was probably facilitated by minor irregularities in the lens surfaces and by the enhanced van der Waals attraction between the two halite lenses. The subsequent growth of the halite grain boundary would then act to exclude any surface-adsorbed fluid layers that might have been present. In contrast, aggressive wetting of the contacts in the halite/silica experiments by brine (i.e. $\theta \approx 0°$) is suggested by the absence of dendritic structures or fluid inclusions within the contact spots and by absence of air at the contact between the two lenses fabricated dry at 90° C (Figure 5). Thus, in the halite/silica experiments, the driving force for expulsion of an adsorbed fluid layer between the halite and fused silica lenses would have been negligible, allowing a high-diffusivity interphase boundary to persist throughout the experiments.

***Transport Properties Within Thin Films***: Electrical double-layer interactions within a thin intergranular fluid film can, in principle, lead to a decrease in diffusion coefficients of several orders of magnitude over that in a bulk-fluid [*Rutter*, 1976]. But, in light of the small Debye lengths calculated, we expect the contribution of the electroviscous effect to transport along the interphase boundary film to be small. Further, we suppose that this film may be maintained largely by repulsive hydration forces. Although a quantitative description of factors controlling the diffusivity ($D_b\delta$) of hydrated intergranular films is not available, *Tada et al.* [1987] proposed that the diffusivity of such a film might be considerably less than that calculated from the electroviscous effect [*Rutter*, 1976], and that the diffusivity probably decreases with decreasing film thickness. If so, then the increase in surface separation with decreasing normal stress anticipated for such a film could lead to a coupled increase in $D_b$ and might explain the apparent non-linearity in normalized convergence rates from our halite/silica experiments at 50° C and $P_m$<1 MPa (Figure 19). Alternatively, this departure from linearity might result from a change in boundary structure at very low normal stresses.

We can not explain the absence of a significant temperature effect in the halite/silica experiments in brine. Presumably, the temperature invariance must reflect changes in either the structure or the thickness of the interphase-boundary film sufficient to offset the expected thermal activation. In this regard, measurements of the disjoining pressure between fused silica plates in distilled water show a rapid thinning of the intervening fluid layer with increasing temperature for $\delta \leq 10$ nm [*Peschel and Adlfinger*, 1971]. Those results indicate that plate separation at $P_m = 0.1$ MPa decreases from about 12.5 nm at 10° C to 3.7 nm at 50° C and then drops to zero at temperatures above 74° C. Similarly, *Derjaguin et al.* [1978] observed that repulsive hydration forces between silica plates in water disappear at temperatures greater than about 65° C. If $D_b$ and $\delta$ are positively correlated, as suggested by *Tada et al.* [1987; see also *Rutter*, 1976], then a small decrease in $\delta$ with increasing temperature in our experiments might be sufficient to counteract the increase in rates calculated above by assuming bulk electrolyte properties and a constant $\delta$.

## Effect of Thin Clay Films on the Kinetics of Pressure Solution

The halite/clay/silica experiment (PSH 27) was, in a sense, an extreme example of the effects of interphase boundaries, because a large number of (presumably hydrated) interfaces were combined into a single layer. The convergence rate in the clay experiment was about five times greater than that in the halite/silica experiments at the same load (Figure 17), providing direct experimental confirmation of the suggestion by a number of authors that clays, especially illites and

smectites, promote IPS and the localized deformation along stylolites and tectonic cleavage [e.g. *Weyl*, 1959; *Engelder and Marshak*, 1985; *Houseknecht and Hathon*, 1987; *Tada and Siever*, 1989].

The rate-controlling step for IPS in the halite/clay/silica experiment is unknown. If, however, we assume (by analogy to the halite/silica experiments) that convergence in the clay experiment was rate-limited by diffusion through the interphase boundary layer, we can use Equation 5 to estimate $D_b\delta$ for the clay layer. Using $\overline{P}_m$, $\overline{a}^2$ and $\Delta s/\Delta t$ for this experiment over elapsed times of 129–359 hours yields $D_b\delta \approx 3.7 \times 10^{-11}$ cm$^3$/s, about 30 times greater than the highest $D_b\delta$ estimated from the halite/silica experiments at 50° C. If $\delta \approx 0.8$ μm for the montmorillonite layer, then the apparent (or bulk) diffusion coefficient for this layer is about $4.6 \times 10^{-7}$ cm$^2$/s, about in the middle of the range of diffusion coefficients estimated for our halite/silica experiments at the same temperature ($1.3 \times 10^{-5}$ to $6.1 \times 10^{-8}$ cm$^2$/s). If the diffusion-limited model is appropriate to this experiment, then montmorillonite can maintain a high volume percentage of adsorbed (structured) water or interparticle bulk water between grains even while under load. In fact, such an interpretation agrees with simultaneous measurements of swelling pressure and c-axis spacing in a Na-montmorillonite immersed in a $10^{-4}$ N NaCl solution [*Viani et al.*, 1983]. Under an applied normal stress of 0.8 MPa (comparable to $\overline{P}_m$ during our halite/clay/silica experiment) the thickness of the interlayer water in the Rio Escondido montmorillonite was about three times greater than that ot the intervening silicate layers (i.e. a water/clay ratio of about 3/1).

The ability of montmorillonite to adsorb interlayer and surface water and swell even while under load made it an attractive candidate for experiments. For this reason we must consider the possibility that clay extruded from under the contact. We can rule this possibility out, however, because extrusion of clay would have caused the clay coating at the perimeter of the contact to warp or buckle. As described above, we used a monochrometer to verify that the Newton's rings arose from reflections from the upper surface of the clay layer and the lower surface of the halite lens. Thus any distortion or thickening within the clay layer adjacent to the contact spot would have been clearly visible as perturbations to the lower-order interference fringes. These perturbations were not observed (Figure 15).

**Relevance to Fault-Zone Rheology:**

Even in the absence of pressure solution, gradients in surface curvature give rise to two other types of solution-transport processes: neck growth and crack healing. During crack healing, material diffuses from the flat wall of a crack to the highly curved crack tip, resulting in infilling of the crack with the host material [e.g. *Brantley et al.*, 1989]. During neck growth, material is dissolved along the wall of an open pore, diffuses through the pore fluid and precipitates at the sharply curved contacts between grains [*Hickman and Evans*, 1992].

Pressure solution, neck growth and crack healing can affect fault-zone rheology in markedly different ways. Pressure solution leads to time-dependent convergence or slip between grains or asperities in contact, resulting in densification and/or shear creep of a fault gouge or sediment. Neck growth, however, leads to an increase in the size of the contact between two grains or asperities, resulting in induration and strengthening. Theoretical analysis has shown that pressure solution might result in aseismic creep along faults and a reduction in overall fault strength [*Rutter and Mainprice,* 1979]. Neck growth and crack healing, however, may lead to the time-dependent strengthening of faults between earthquakes [e.g. *Fredrich and Evans*, 1992] or allow the formation of pressure seals along faults which prevent the escape of overpressured pore fluids [e.g. *Sleep and Blanpied*, 1992; *Byerlee*, 1993]. Which of these processes dominates along any given fault will depend on a variety of factors, including temperature, stress, fluid chemistry, pore or crack geometry, grain size and mineralogy.

Mineralogic observations on gouges from the San Andreas and other mature fault zones at shallow depths indicate that these gouges are mineralogically diverse, contain large percentages of illites, smectites and other clays, and are very fine-grained [e.g. *Wu*, 1978; *Chester et al.*, 1993]. In the experiments reported here, we observed no neck growth and a dramatic acceleration of pressure-solution rates at the interfaces between dissimilar materials (halite and fused silica) and in the presence of intergranular clay. These results suggest that intergranular clays and the

juxtaposition of different minerals in the Earth can poison neck growth and, by enhancing interphase boundary diffusivities, lead to low-strain-rate weakening at mid- to upper-crustal levels within active fault zones. Also, the inverse-square dependance of pressure solution rates on contact spot size observed in our experiments, if applicable to other mineral systems, suggests that pressure solution creep in polycrystalline aggregates should follow an inverse cube dependance on grain size [*Rutter*, 1976], further accelerating IPS rates in fine grained fault gouges.

## CONCLUSIONS

Deformation in the contact region between the lenses occurred in two stages. During the initial loading, the halite lens flattened essentially instantaneously. After this stage, the lenses converged at a rate which depended on both the load and the contact area. Etch pit studies, observations of plastic surface flattening, and measurements of mean contact stress as a function of load indicate that plastic yielding occurred during the initial loading in all experiments. In spite of high dislocation densities in the contact zone, no metastable, island-channel boundaries were observed. After the initial plastic flattening, the convergence rates between polished halite and silica lenses in brine at loads of 0.11–4.22 N and at temperatures of 8.3°–90.2° C ranged from 0.01 to 0.06 μm/day. These rates were much greater than the convergence between two halite lenses under similar conditions [*Hickman and Evans, 1991*]. Creep experiments conducted in dry nitrogen and moist air under the most extreme conditions of load and temperature indicate that dislocation creep did not contribute to the observed convergence. As no undercutting or cataclasis was observed, this convergence must be occurring by diffusion through a high-diffusivity, intergranular fluid film.

Measurements of convergence rates between halite and silica lenses in brine at 50.2° C, and loads of 0.11–4.22 N, indicate that convergence in this system is rate-limited by interphase-boundary diffusion rather than dissolution or precipitation. The product of the boundary diffusion coefficient, $D_b$, and the boundary thickness, $\delta$, ranges from $1.3 \times 10^{-12}$ $cm^3/s$ to $1.8 \times 10^{-13}$ $cm^3/s$ at mean stresses of 0.56–7.9 MPa. Although $D_b\delta$ appears to be relatively constant over most of the range of normal stresses investigated, it may increase at normal stresses less than about 0.5 MPa. If $\delta$ is between 1–30 nm, $D_b$ must range between $1.3 \times 10^{-5}$ to $6.1 \times 10^{-8}$ $cm^2/s$. This diffusion coefficient is many orders of magnitude greater than $D_b$ for solid-state grain boundaries at higher temperatures, but is comparable to, or somewhat less than, the mutual diffusion coefficient of NaCl in bulk brine at 50° C.

Halite/silica experiments conducted in brine at a constant load of 0.11 N, but at temperatures of 8.3°, 50.2° and 90.2° C, indicate virtually no dependence of convergence rates upon temperature. This result contrasts with the 4 to 5 fold increase in rates predicted assuming a constant boundary thickness, $\delta$, and an activation energy for transport within the intergranular fluid film equal to that for transport through bulk brine. Although the mechanisms of diffusion through structured fluid layers are poorly understood, the near-zero activation energies observed may reflect changes in interphase boundary structure or thickness with increasing temperature that are sufficient to offset thermal activation. Experiments using other minerals will be necessary before we can determine if temperature insensitivity is a general feature of IPS in water films or if this behavior is simply a peculiarity of the halite/silica system.

Thin clay films result in dramatic accelerations of pressure solution. An experiment using halite and silica lenses with an intervening montmorillonite layer 0.8-μm-thick indicates a five-fold increase in convergence rates over that observed in halite/silica experiments in brine, but with no clay layer. The rapid convergence in the halite/clay/silica experiment agrees with geological evidence showing that intergranular clays and clay seams localize solution transfer deformation. Although the rate-limiting step for convergence in the halite/clay/silica experiment is not known, assuming that convergence was limited by interphase boundary diffusion, we estimate that $D_b\delta$ is about 30 times greater than in the halite/silica experiments.

The accelerating effects of dissimilar interfaces and clay films, as well as the inverse square relationship between convergence rate and contact spot size, suggests that fine-grained, mixed-phase aggregates in the earth should be particularly susceptible to pressure-solution deformation.

In particular, fine grain sizes and high clay contents could lead to pronounced weakening at low strain-rates, and to fault creep at mid- to upper-crustal levels. Our results also suggest that even a small percentage of clays or other second phases in sedimentary deposits could dramatically enhance the rates of densification and diagenesis.

## REFERENCES


Adamson, A. W., Physical Chemistry of Surfaces, fourth edition, John Wiley and Sons, New York, 664 pp., 1982.

Angelvine, C. L., D. L. Turcotte, and M. D. Furnish, Pressure solution lithification as a mechanism for the stick-slip behavior of faults, *Tectonics, 1*, 151-160, 1982.

Bathurst, R. C. G., Diagenetic fabrics in some British Dinantian limestones, *Liverpool Manchester Geol. J., 2*, 11-36, 1958.

Bockris, J. O'M., and A. K. N. Reddy, *Modern Electrochemistry*, 2 vol, Plenum, New York, 1432 pp., 1970.

Bosworth, W., Strain-induced preferential dissolution of halite, *Tectonophysics, 78*, 509-525, 1981.

Brantley, S. L., B. Evans, S.H. Hickman, and D.A. Crerar, Healing of microcracks in quartz: Implication for fluid flow, *Geology,18,* 136-139, 1989.

Byerlee, J.D., Friction, overpressure and fault normal compression, *Geophys. Res. Lett.,17*, 2109-2112, 1990.

Byerlee, J.D., A model for episodic flow of high pressure water in fault zones before earthquakes, *Geology, 21*, 303-306, 1993.

Carter, N. L., and H. C. Heard, Temperature and rate dependent deformation of halite, *Am. J. Sci., 269,* 193-249, 1970.

Chester, F.M., J.P. Evans and R.L. Biegel, Internal structure and weakening mechanisms of the San Andreas fault, *J. Geophys. Res., 98*, 771-786, 1993.

Chester, F. M., & Higgs, N. G., Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions, *J. Geophys. Res.*, **97**, 1859-1870,1992.

Chester, F. M., and J. M. Logan, Implications for mechanical properties of brittle faults from observations of the Punchbowl fault zone, California, *Pure Appl. Geophys., 124*, 79-106, 1986.

Clarke, D. R., On the equilibrium thickness of intergranular glass phases in ceramic materials, *J. Am. Ceram. Soc., 70*, 15-22, 1987a.

Clarke, D. R., Grain boundaries in polycrystalline ceramics, *Ann. Rev. Mater. Sci., 17*, 57-74, 1987b.

Davidge, R. W., and P. L. Pratt, Plastic deformation and work- hardening in NaCl, *Phys. Stat. Sol., 6*, 759-776, 1964.

Davis, L. A., and R. B. Gordon, Pressure dependence of the plastic flow stress of alkali halide single crystals, *J. Appl. Phys., 39,* 3885-3897, 1968.

de Boer, R. B., On the thermodynamics of pressure solution - interaction between chemical and mechanical forces, *Geochim. Cosmochim. Acta, 41,* 249-256, 1977.

de Boer, R. B., P. J. C. Nagtegaal, and E. M. Duyvis, Pressure solution experiments on quartz sand, *Geochim. Cosmochim. Acta, 41,* 257-264, 1977.

Dennis, S. M., and B. K. Atkinson, The influence of water on the stress supported by experimentally faulted Westerly granite, *Geophys. J. Roy. Astr. Soc., 71*, 285-294, 1982.

Derjaguin, B. V., Z. M. Zorin, N. V. Churaev, and V. A. Shishin, Examination of thin layers of liquids on various solid substrates, in: J. F. Padday (ed.), *Wetting, Spreading and Adhesion,* Academic Press, London, 201-236, 1978.

Durnay, D. W., and J. G. Ramsay, Incremental strains measured by syntectonic crystal growths, in *Gravity and Tectonics,* K.A. Dejong and R. Schoften (eds.), pp. 67-97, John Wiley & Sons, 1973.

Elliot, D., Diffusion flow laws in metamorphic rocks, *Geol. Soc. Amer. Bull., 84*, 2645-2664, 1973.

Engelder, T., and S. Marshak, Disjunctive cleavage formed at shallow depths in sedimentary rocks, *J. Struct. Geol., 7*, 327-343, 1985.

ESCO Products Inc., *Optical Materials and Components Handbook*, Esco Products Inc., Oak Ridge, New Jersey, 37 pp., 1986.

Etzler, F. M., and W. Drost-Hansen, Recent thermodynamic data on vicinal water and a model for their interpretation, *Croatica Chemica Acta, 56*, 563-592, 1983.
Fredrich, J. and B. Evans, Strength recovery along simulated faults by solution transfer processes, *33rd U.S. Symposium on Rock Mechanics*, 121-130, 1992.
Gibbs, R. J., Error due to segregation in quantitative clay mineral X-ray diffraction mounting techniques, *Am. Mineral., 50*, 741-751, 1965.
Gratier, J.-P., Experimental pressure solution of halite by an indenter technique, *Geophys. Res. Lett., 20*. 1647-1650, 1993.
Gratier, J. P., and R. Guiguet, Experimental pressure solution-deposition on guartz grains: the crucial effect of the nature of the fluid, *J. Struct. Geol., 8*, 845-856, 1986.
Green, H. W., "Pressure solution" creep: some causes and consequences, *J. Geophys. Res., 89*, 4313-4318, 1984.
Hanneman, R. E., and J. H. Westbrook, Effects of adsorption on the indentation deformation of non-metallic solids, *Phil. Mag., 18*, 73-88, 1968.
Hickman, S., Experimental studies of pressure solution and crack healing in halite and calcite, Ph.D. dissertation, Massachusetts Institute of Technology, 176 pp., 1989.
Hickman, S., and B. Evans, Experimental pressure solution in halite: the effect of grain/interphase boundary structure, *J. Geol. Soc. London, 148*, 549-560, 1991.
Hickman, S., and B. Evans, Growth of grain contacts in halite by solution transfer: Implications for diagenesis, lithification, and strength recovery, in *Fault Mechanics and Transport Properties of Rocks*, B. Evans and T.-F. Wong (eds.), Academic Press, pp. 253-280, 1992.
Hiemenz, P. C., *Principles of Colloid and Surface Chemistry*, second ed., Marcel Dekker, New York, 815 pp., 1986.
Houseknecht, D. W., and L. A. Hathon, Petrographic constraints on models of intergranular pressure solution in quartzose sandstones, *Applied Geochemistry, 2*, 507-521, 1987.
Israelachvili, J. N., *Intermolecular and Surface Forces, with Applications to Colloidal and Biological Systems*, Academic Press, London, 296 pp., 1985a.
Israelachvili, J. N., Measurements of hydration forces between macroscopic surfaces, *Chemica Scripta*, 25, 7-14, 1985b.
Johnson, K. L., *Contact Mechanics*, Cambridge University Press, Cambridge, 452 pp., 1987.
Kirby, S. H., Tectonic stresses in the lithosphere: constraints provided by the experimental deformation of rocks, *J. Geophys. Res., 85*, 6353-6363, 1980.
Latanision, R. M., Surface effects in crystal plasticity: general overview, in: R. M. Latanision and J. T. Fourie (eds.), *Surface Effects in Crystal Plasticity*, Nato Advanced Study Institute Series E, Noordhoff, Leyden, No. 17, 1977.
Lehner, F. K., Thermodynamics of rock deformation by pressure solution, in: D. Barber and P. G. Meridith (eds.), *Deformation Processes in Minerals, Ceramics, and Rocks*, Unwin Hyman, London, 296-333, 1990.
Matthews, J.R., Indentation hardness and hot pressing, *Acta. Metallurgica, 28*, 311-318, 1980.
Mendelson, S., Dislocation etch pit formation in sodium chloride, *J. Appl. Phys., 32*, 1579-1583, 1961.
Olander, D. R., A. J. Machiels, M. Balooch, and S. K. Yagnik, Thermal gradient migration of brine inclusions in synthetic alkalai halide single crystals, *J. Appl. Phys., 53*, 669-681, 1982.
Pashley, R. M., DLVO and hydration forces between mica surfaces in $Li^+$, $Na^+$, $K^+$, and $Cs^+$ electrolyte solutions: A correlation of double-layer and hydration forces with surface cation exchange properties, *J. Colloid Interface Sci., 83*, 531-546, 1981.
Pashley, R. M., The effects of hydrated cation adsorption on surface forces between mica crystals and its relevance to colloidal systems, *Chemica Scripta, 25*, 22-27, 1985.
Pashley, R. M., and J. N. Israelachvili, Molecular layering of water in thin films between mica surfaces and its relation to hydration forces, *J. Colloid Interface Sci., 101*, 511-523, 1984.
Paterson, M. S., Nonhydrostatic thermodynamics and its geologic applications, *Rev. Geophys. Space Phys., 11*, 355-389, 1973.
Peschel, G., and K. H. Aldfinger, Temperaturabhangigkeit der Viskositat sehr dunner Wasserschichten zwischen Quarzglasoberflachen, *Naturwissenschaften, 56*, 558, 1969.

Peschel, G., and K. H. Aldfinger, Thermodynamic investigations of thin liquid layers between solid surfaces II. Water between entirely hydroxylated fused silica surfaces, *Z. Naturforschung, 26A,* 707-715, 1971.
Peschel, G., P. Belouschek, M. M. Muller, M. R. Muller, and R. Konig, The interaction of solid surfaces in aqueous systems, *Colloid Polymer Sci., 260,* 444-451, 1982.
Pharr, G. M., and M. F. Ashby, On creep enhanced by a liquid phase, *Acta Mettalurgica., 31,* 129-138, 1983.
Power, W. L., and T. E. Tullis, The relationship between slickenside surfaces in fine-grained quartz and the seismic cycle, *J. Struct. Geol., 11,* 879-893, 1989.
Raj, R., Creep in polycrystalline aggregates by matter transport through a liquid phase, *J. Geophys. Res., 87,* 4731-4739, 1982.
Raj, R., and C. K. Chyung, Solution-precipitation creep in glass ceramics, *Acta Metall., 29,* 159-166, 1981.
Rice, J. R., Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault, in *Fault Mechanics and Transport Properties in Rocks: A Festschrift for W. F. Brace,* (ed. Evans, B. & Wong, T. F.) 475-503 (Academic Press, London, 1992).
Robin, P.-Y. F., Pressure solution at grain-to-grain contacts, *Geochim. Cosmochim. Acta, 42,* 1383-1389, 1978.
Rossi, B., *Optics,* Addison-Wesley, Redding, Mass., 510 pp., 1957.
Rudnicki, J., Physical models of earthquake instability and precursory processes, *Pure Appl. Geophys.* **126,** 531-524, 1988.
Rutter, E. H., The kinetics of rock deformation by pressure solution, *Phil. Trans. Roy. Soc. London, A, 283,* 203-219, 1976.
Rutter, E. H., Pressure solution in nature, theory and experiment, *J. Geol. Soc. London, 140,* 725-740, 1983.
Rutter, E. H., and D. H. Mainprice, The effect of water on stress relaxation of faulted and unfaulted sandstone, *Pure Appl, Geophys, 116,* 634-654, 1978.
Rutter, E. H., and D. H. Mainprice, On the possibility of slow fault slip controlled by a diffusive mass transfer process, *Gerlands Beitr. Geophys., 88,* 154-162, 1979.
Shewmon, P. G., *Diffusion in Solids,* J. Williams Book Co., 203 pp., 1983.
Sibson, R. H., Continental fault structure and the shallow earthquake source, *J. Geol. Soc. London, 140,* 741-767, 1983.
Simmons, G., and H. Wang, *Single Crystal Elastic Constants and Calculated Aggregate Properties: A handbook,* 2nd ed., M.I.T. Press, Cambridge, 1971.
Sleep, N. H., and M. L. Blanpied,Creep, compaction and the weak rheology of major faults, *Nature, 359,* 687-692, 1992.
Spiers, C. J., and P. M. T. M. Schutjens, Densification of crystalline aggregates by fluid phase diffusional creep, in: D. Barber and P. G. Meridith (eds.), *Deformation Processes in Minerals, Ceramics, and Rocks,* Unwin Hyman, London, 334-353, 1990.
Spiers, C.J., P.M.T.M. Schutjens, R.H. Brzesowsky, C.J. Peach, J.L. Liezenberg and H.J. Zwart, Experimental determination of constitutive parameters governing creep of rocksalt by pressure solution, In: R.J. Knipe and E.H. Rutter (eds.), *Deformation Mechanisms, Rheology and Tectonics,* Geol. Soc. London. Spec. Publ. No. 54, 215-227, 1990.
Sprunt, E. S., and A. Nur, Experimental study of the effects of stress on solution rate, *J. Geophys. Res., 82,* 3013-3022, 1977.
Tada, R., R. Maliva, and R. Siever, A new mechanism for pressure solution in porous quartzose sandstone, *Geochim. et Cosmochim. Acta, 51,* 2295-2301, 1987.
Tada, R., and R. Siever, Experimental knife-edge pressure solution of halite, *Geochim. et Cosmochim. Acta, 50,* 29-36, 1986.
Tada, R., and R. Siever, Pressure solution during diagenesis, *Ann. Rev. Earth Planet. Sci., 17,* 89-118, 1989.
Timoshenko, S. P., and J. N. Goodier, *Theory of Elasticity,* 3rd ed., McGraw-Hill, New York, 567 pp., 1970.
Trurnit, P., Pressure solution phenomena in detrital rocks, *Sediment. Geol., 2,* 89-114, 1968.

Urai, J. L., C. J. Spiers, H. J. Zwart, G. S. Lister, Weakening of rock salt by water during long-term creep, *Nature, 324,* 554-557, 1986.
Viani. B. E., P. F. Low, and C. B. Roth, Direct measurements of the relation between interlayer force and interlayer distance in the swelling of montmorillonite, *J. Colloid Interface Sci., 96,* 229-244, 1983.
Vitagliano, V., Determinazione delle mobilita ioniche per le soluzioni acquose di NaCl a diverse temperature, *Gazzetta Chimica Italiana, 90,* 1847-1858, 1960.
Westbrook, J. H., and P. J. Jorgensen, Indentation creep of solids, *Trans. Met. Soc. of AIME, 233,* 425-428, 1965.
Weyl, P. K., Pressure solution and the force of crystallization - a phenomenological theory, *J. Geophys. Res., 64,* 2001-2025, 1959.
Wiederhorn, S. M., and E. R. Fuller, Jr., Effect of surface forces on subcritical crack growth in glass, *J. Am. Ceram. Soc., 72,* 248-251, 1989.
Wong, T-F., Mechanical compaction and brittle-ductile transition in porous sandstones. *In*: Knipe, R. J. & Rutter, E. H. (eds) *Deformation Mechanisms, Rheology, and Tectonics*, Geological Society of London, Special Publication, 54, 111-122, 1990.
WU, F. T., Mineralogy and physical nature of clay gouge, *Pure Appl. Geophys., 116,* 655-689, 1978.

## TABLE 1: PRESSURE SOLUTION EXPERIMENTS

| Experiment | Temperature[A] (°C) | Radius of Curvature[B] (cm) | Fluid | Load (Newtons) | Fabri-cation[C] | Experiment Duration (Hours) | Weight Change (Percent) | How Convergence Measured[D] |
|---|---|---|---|---|---|---|---|---|
| *Halite/Silica Experiments:* | | | | | | | | |
| PSH 17 | 50.19±0.20 | 9.29 | brine | 4.22±0.11 | wet | 290.3 | +0.2 | a |
| PSH 18 | 50.19±0.05 | 9.16 | nitrogen | 4.21±0.11 | dry | 236.1 | 0.0 | a |
| PSH 22 | 50.20±0.09 | 8.79 | brine | 0.762±0.023 | wet | 572.0 | +0.2 | a |
| PSH 23 | 50.21±0.07 | 7.10 | brine | 0.111±0.004 | wet | 527.9 | +0.2 | a |
| PSH 26 | 50.23±0.08 | 8.58 | brine | 0.763±0.023 | wet | 572.6[E] | -0.1 | a |
| PSH 28 | 90.05±0.19 | 7.82 | brine | 0.111±0.005 | wet[F] | 179.2 | +1.0 | a |
| PSH 29 | 90.16±0.17 | 7.17 | brine | 0.111±0.005 | dry | 336.8 | -----[G] | a |
| PSH 30 | 90.45±0.21 | 7.95 | nitrogen | 0.113±0.005 | dry | 334.1 | 0.0 | a |
| PSH 31 | 8.29±0.16[H] | 8.18 | brine | 0.114±0.005 | wet | 500.4 | 0.0 | a |
| PSH 34[I] | 23±2 | 7.95 | air | 4.23±0.11 | dry | 187.6 | ----- | r |
| " | 50.18±0.04 | " | " | 4.16±0.11 | ----- | 286.0 | ----- | r |
| " | 89.93±0.15 | " | " | 4.07±0.11 | ----- | 102.7 | 0.0 | r |
| PSH 35 | 20±1 | 8.0-9.1 | air | 0.098-10.79 | ----- | ----- | ----- | ----- |
| *Halite/Montmorillonite/Silica Experiment:* | | | | | | | | |
| PSH 27 | 50.22±0.08 | 8.18 | brine | 0.761±0.023 | dry | 359.0 | +0.3 | r |

A Mean values as recorded by precision thermometer in fluid return line; uncertainties are maximum excursions recorded by thermistor

B For the plano-convex (halite) lens

C Dry: Lid to microscope stage closed before brine injected using brine inlet tubes
Wet: Brine added directly to contact before sample assembled and microscope stage lid closed

D r: Ray method; a: Area method

E Lens repositioned after 1.75 hours

F Fabricated dry and then teased with wire inserted through upper brine inlet tube to wet contact

G Distilled water spilled on lens following experiment, therefore weight change during experiment unknown

H Except for power failure that occurred 95 hours into experiment, during which time the stage (thermistor) temperature increased steadily to 13.4° C in 50 min before returning to run temperature.

I Reused halite lens from experiment PSH 30 (radius of curvature given is from experiment PSH 30)

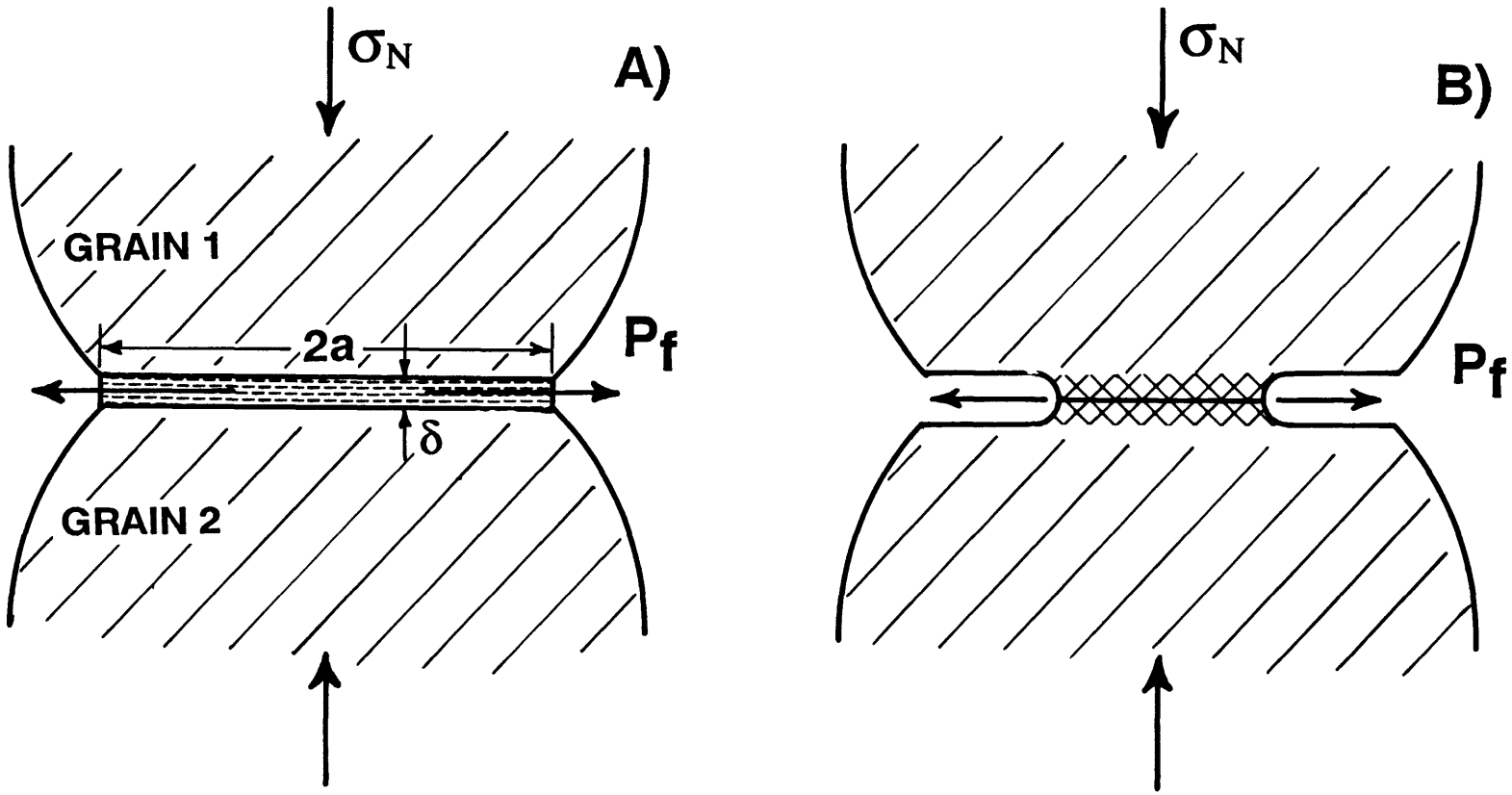



FIGURE 1: Diagram illustrating the two classes of proposed intergranular pressure solution mechanisms: A) water film diffusion and B) undercutting. The normal stress acting on the grain contacts is $\sigma_N$, the pore fluid pressure is $P_f$, the contact spot radius is a, and the thickness of the high diffusivity film is $\delta$.

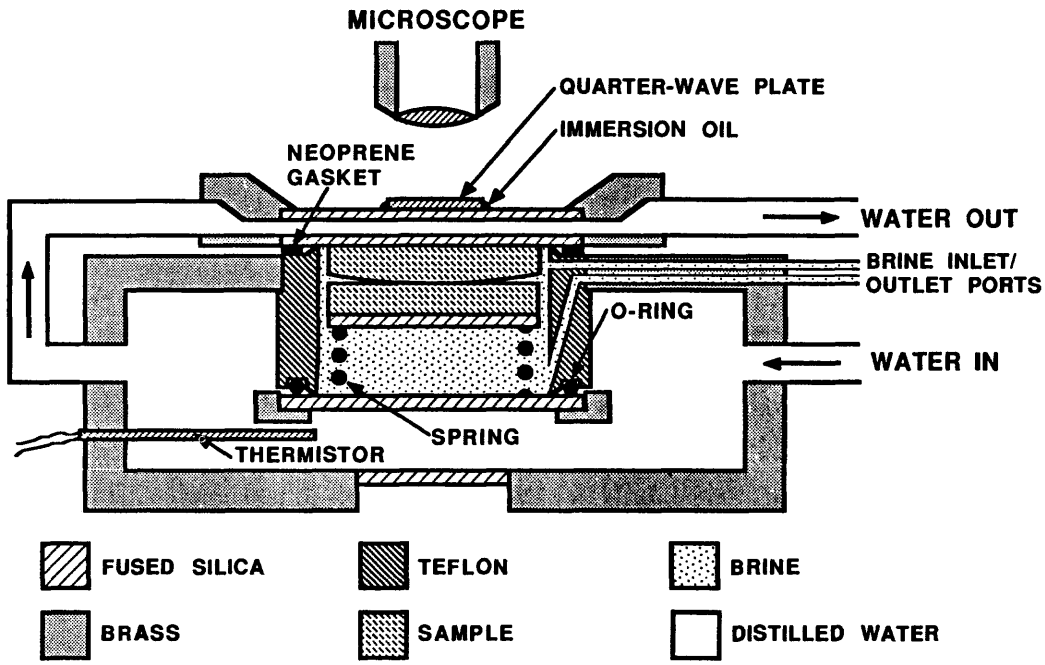



FIGURE 2: Schematic of the heated microscope stage used in this study. The sample column, spring and stagnant brine are contained within a cylindrical teflon and fused-silica chamber which is surrounded on all sides by circulating water from a constant-temperature bath. Loads are adjusted using fused-silica spacers (not shown) beneath the springs. The temperature of the circulating fluid is continuously recorded using a thermistor and measured periodically with a precision thermometer inserted into the water-outlet line. When viewing the sample with reflected sodium light, a quarter-wave plate is mounted on the upper window of the stage and the sample viewed under crossed polars to eliminate reflections from the upper window.

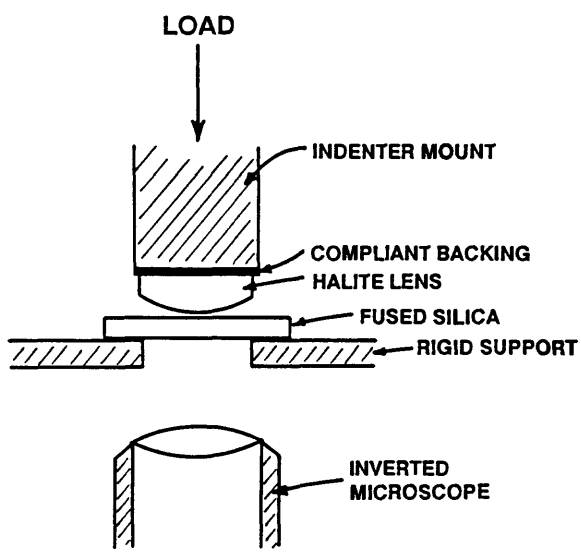



FIGURE 3: Schematic of the modified Tukon Hardness Tester apparatus (manufactured by the Wilson Mechanical Instrument Division of the American Chain and Cable Company, Inc.) used to measure the variation in mean contact stress with load for a halite convex lens pressed against a fused silica flat. The halite lens was mounted on a compliant backing in order to distribute the load evenly across the flat side of the lens.

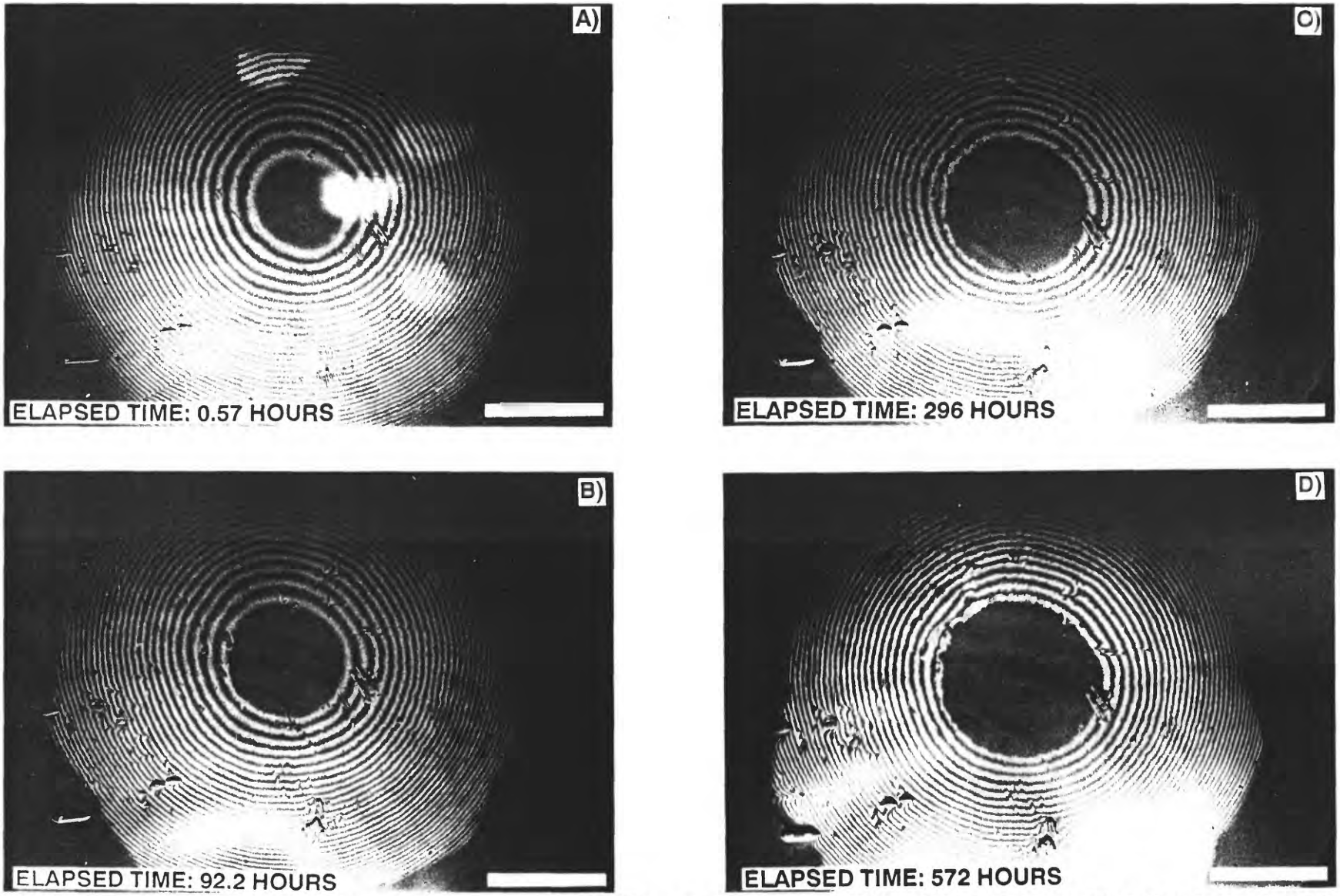



**FIGURE 4:** Reflected sodium light interferograms of a halite/silica experiment (PSH 22) in saturated brine at 50.2° C and a load of 0.76 N for the times shown. The dark concentric bands are interference minima (Newton's rings). As the refractive index of saturated brine is 1.376 at 50.2° C, the change in lens separation between adjacent interference fringes is 0.214 μm. The out-of-focus white patches in A) are the result of air bubbles trapped underneath the halite lens (i.e. not in the plane of the contact spot). The scale bars shown are 0.5 mm.

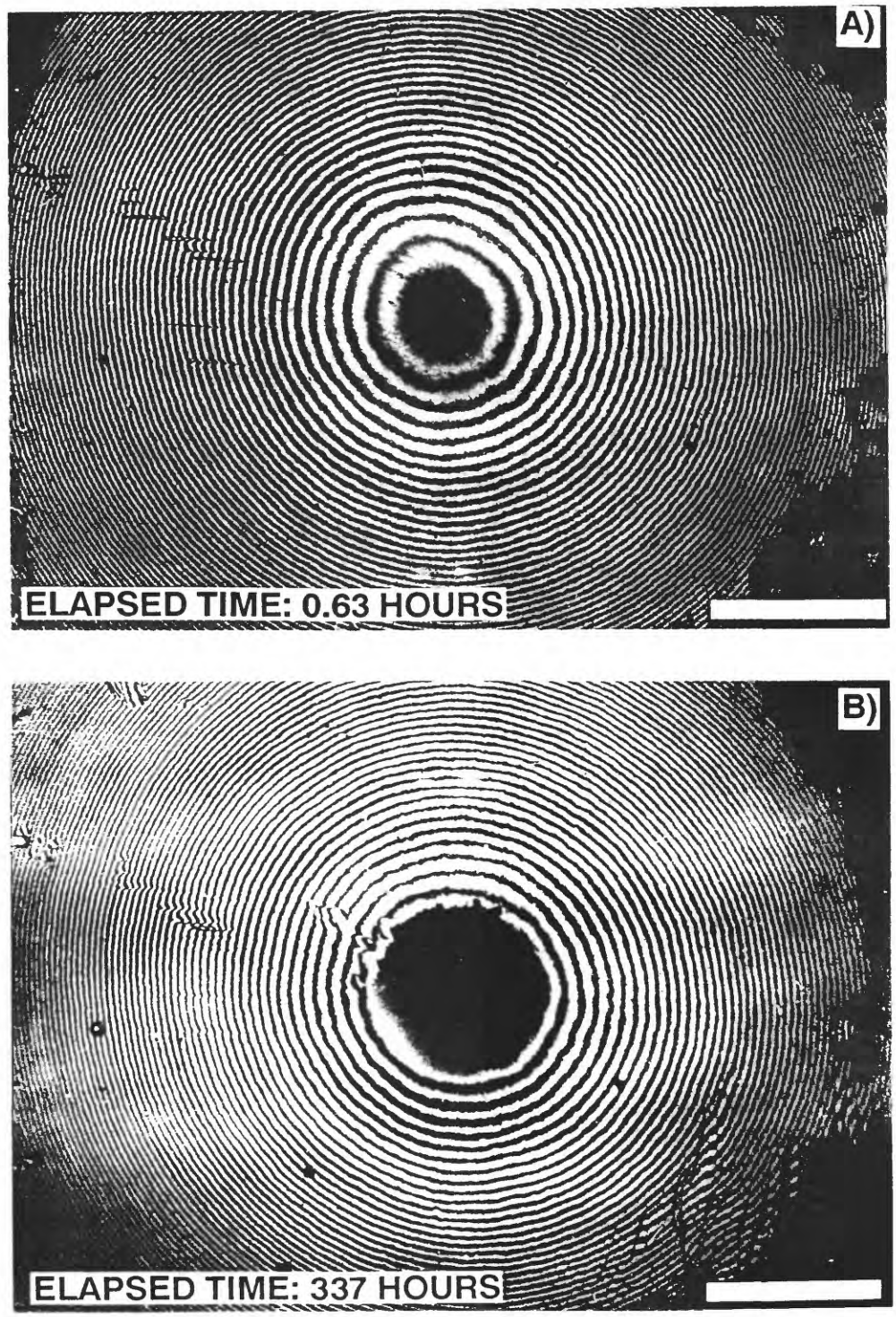



FIGURE 5: Reflected sodium interferogram of a halite/silica experiment (PSH 29) in saturated brine at 90.2° C and 0.11 N for the times shown. Focussing through the sample column shows the light region along the lower left edge of the contact spot in B) to be an interference maxima from the silica/halite interface below the plane of the contact spot. As the refractive index of saturated brine is 1.370 at 90.2° C, the change in lens separation between adjacent interference fringes is 0.215 μm. The scale bars are 0.5 mm.

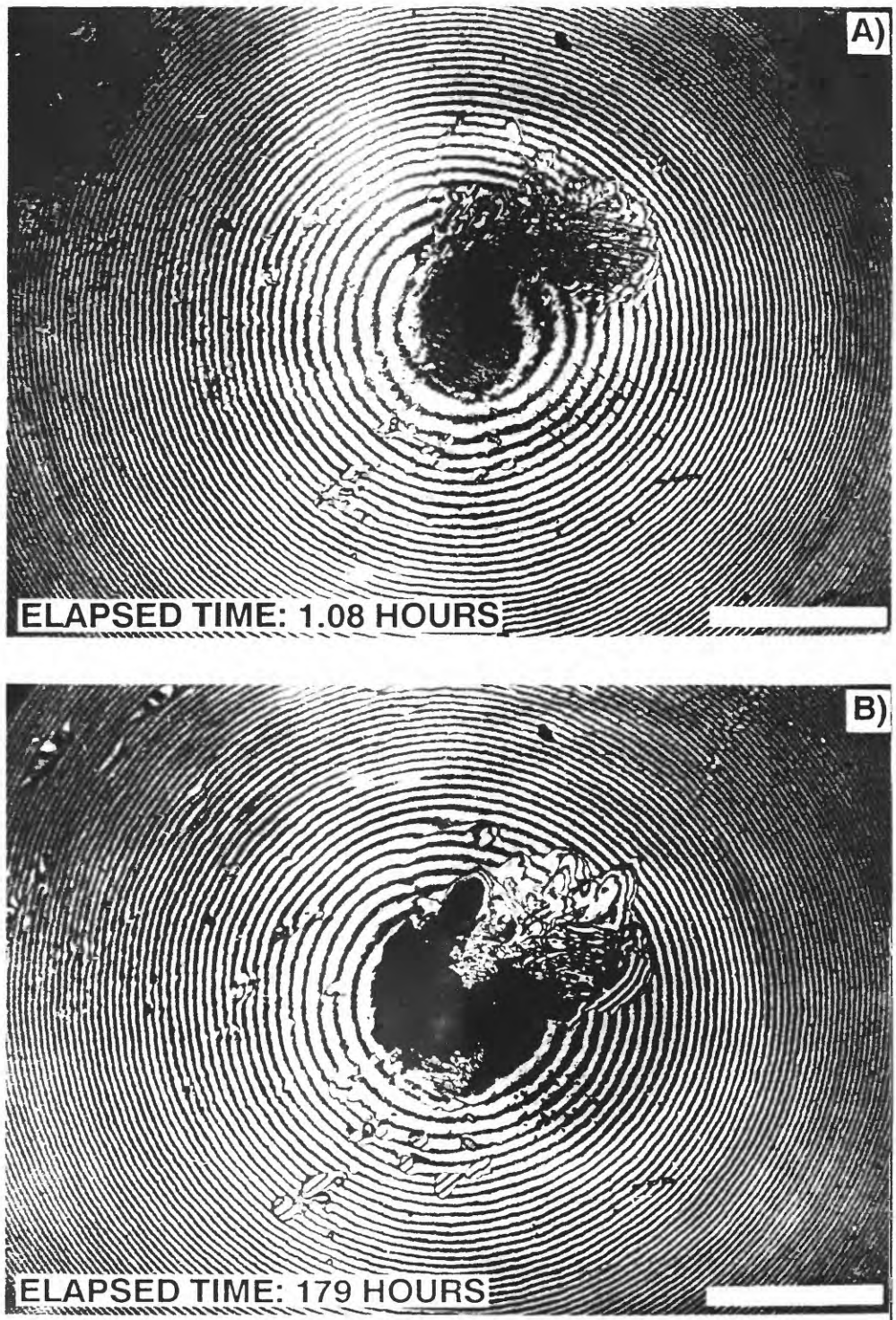



FIGURE 6: Reflected sodium interferograms of a halite/silica experiment (PSH28) in saturated brine at 90.1° C and 0.11 N for the times shown. This sample was fabricated dry and then teased with a thin wire inserted through the upper brine inlet port after the addition of brine to insure wetting of the contact spot. A) Note the extensive damage done to the halite lens during this operation, especially above and to the right of the contact spot, and the presence of numerous fine scratches and open areas within the central contact spot. B) After 179 hours, the halite lens is now in more intimate contact with the silica flat, although the contact spot is still highly irregular. The change in lens separation between adjacent interference fringes is 0.215 μm. The scale bars are 0.5 mm.

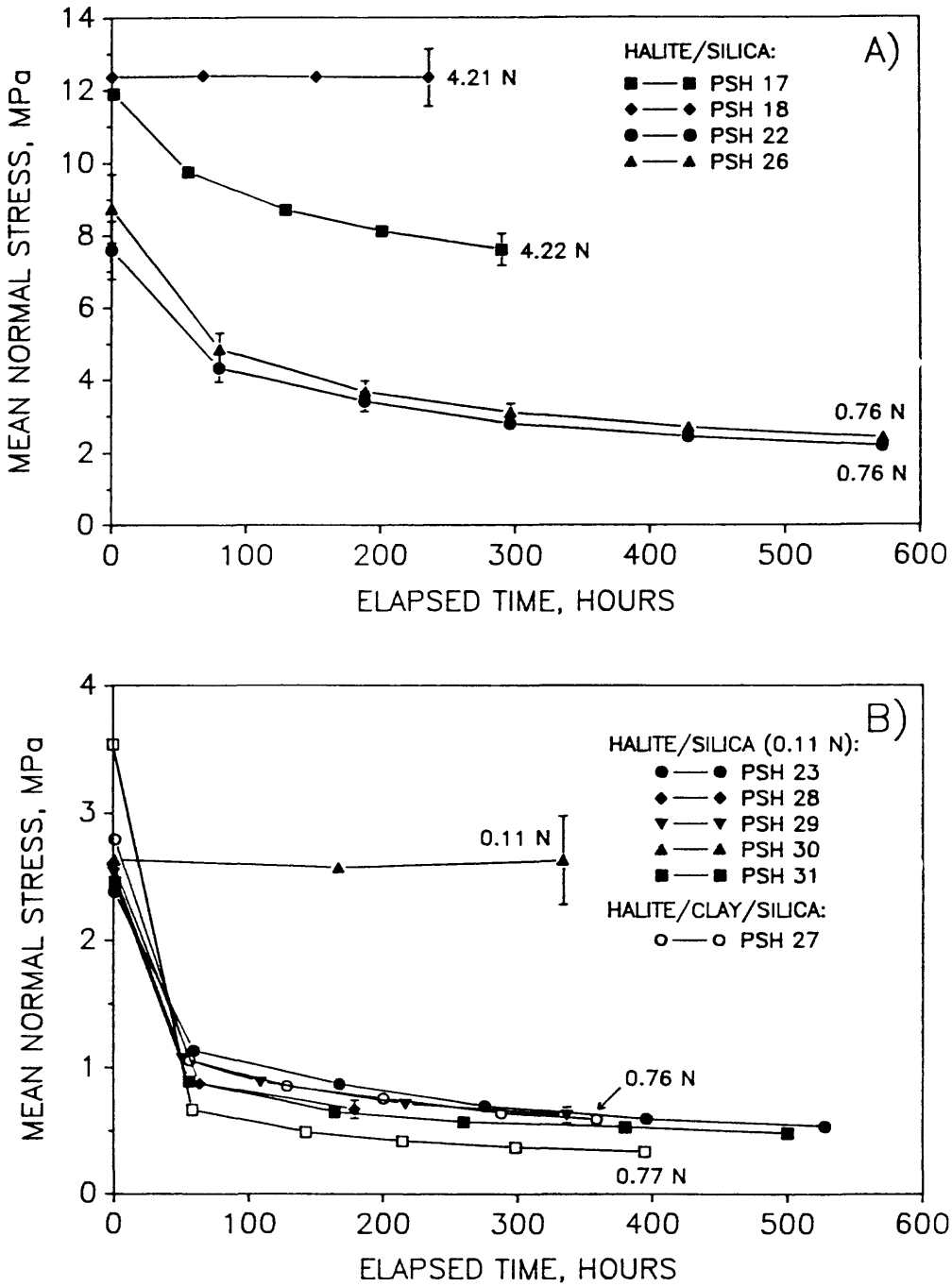



FIGURE 7: Mean effective normal stress, $P_m$, as a function of time for experiments in which a halite lens and silica lens were pressed together (halite/silica) at A) high to intermediate loads and B) low loads. Also shown in B) are contact stresses for an experiment in which a halite convex lens was pressed against a fused silica disk coated with an approximately 0.8-μm-thick montmorillonite layer (halite/clay/silica). With the exception of the halite/silica experiments in B), which were loaded at about 0.11 N (Table 1), the applied loads are indicated for each experiment. Error bars are shown for all data from PSH 22 and PSH 26 to facilitate comparison of data from these experiments conducted under similar conditions (Table 1). Error bars are shown for the last data points in the other experiments, but only when the error is larger than the corresponding symbol.

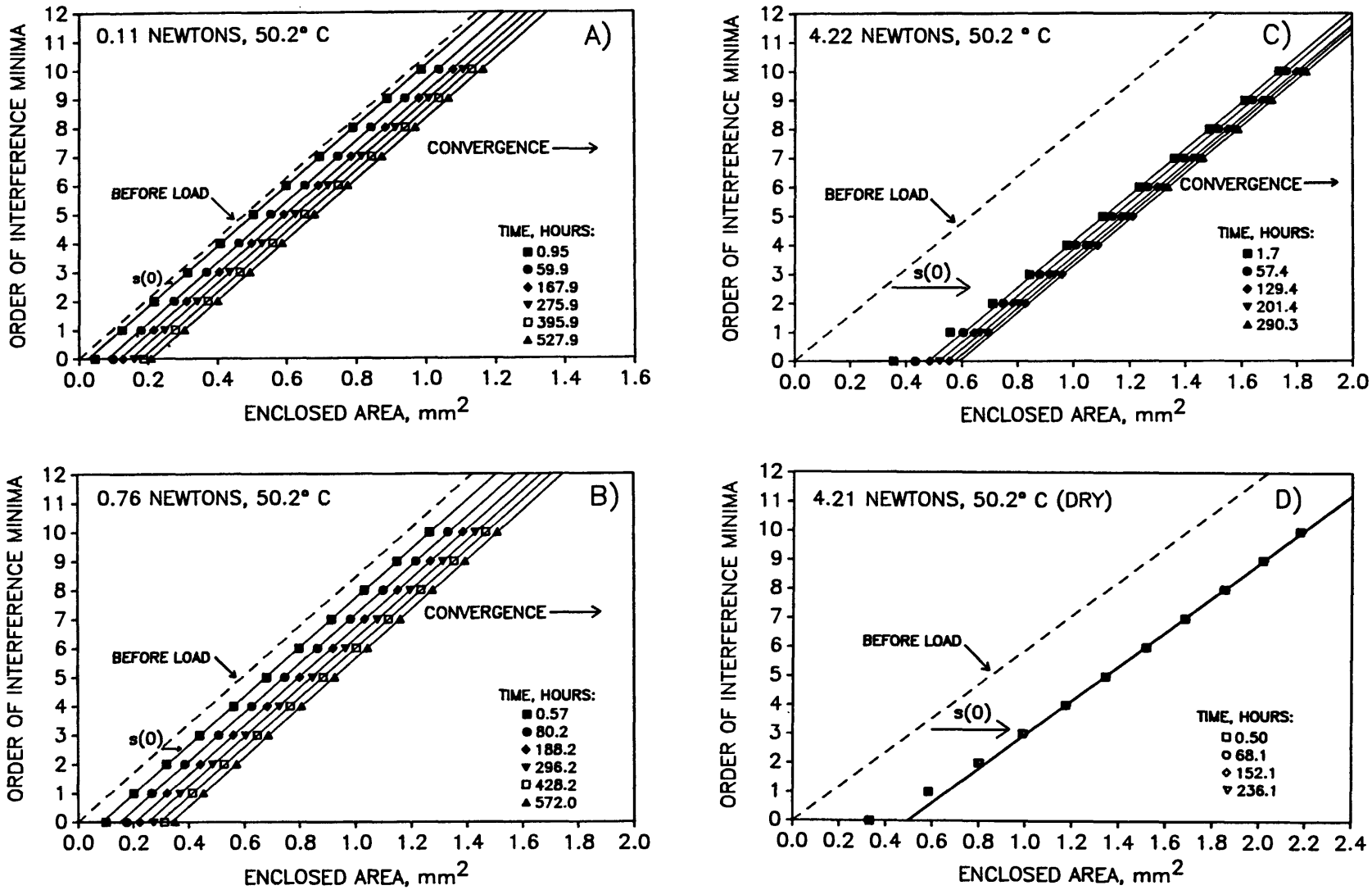



FIGURE 8: Examples of the plots used to measure convergence in halite/silica experiments in brine using Equation 3: A) PSH 23, B) PSH 22 and C) PSH 17. A similar plot is also shown for a halite/silica experiment in dry nitrogen: D) PSH 18. The dashed line indicates the expected relationship between interference minima order number, m, and the enclosed area, $A_m$, before the load is applied. The solid lines shown are linear regressions to the data for m = 1-10 ( A and B); m = 6-10 (C); and m = 3-10 (D). The slopes of these lines in C and D differ because experiment PSH 18 was conducted in dry nitrogen ($n_D$ = 1).

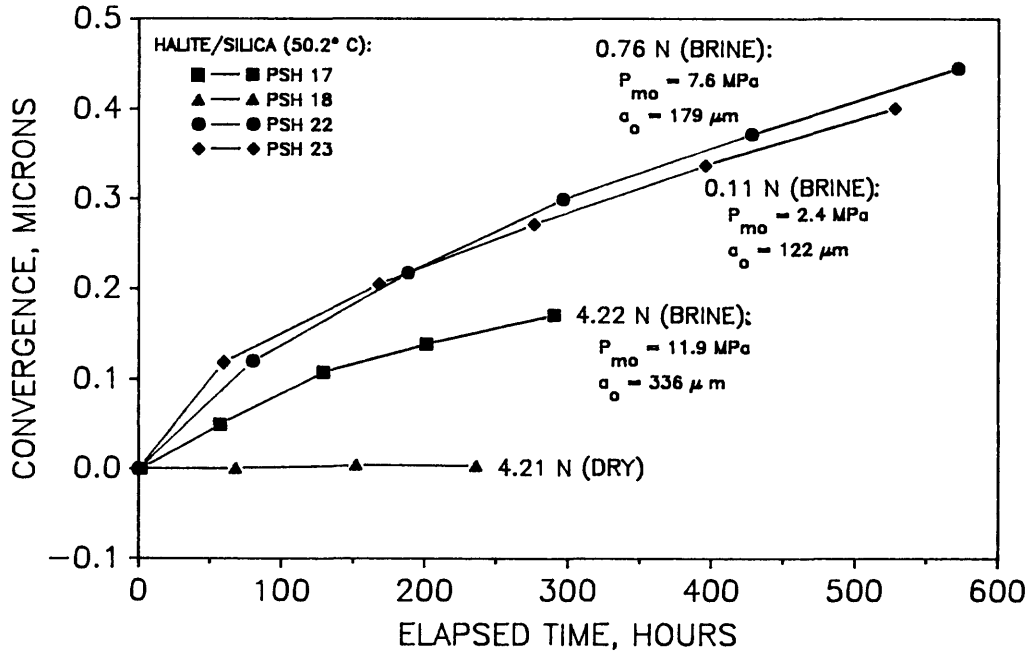



FIGURE 9: Convergence as a function of time for the halite/silica experiments in saturated brine or dry nitrogen at 50.2° C and loads of 0.11-4.22 N. The initial mean effective normal stress, $P_{mo}$, and the initial contact spot radius, $a_o$, are also shown for the experiments in brine. The error bars (at ± one standard deviation) are about the same size as the symbols and are not shown.

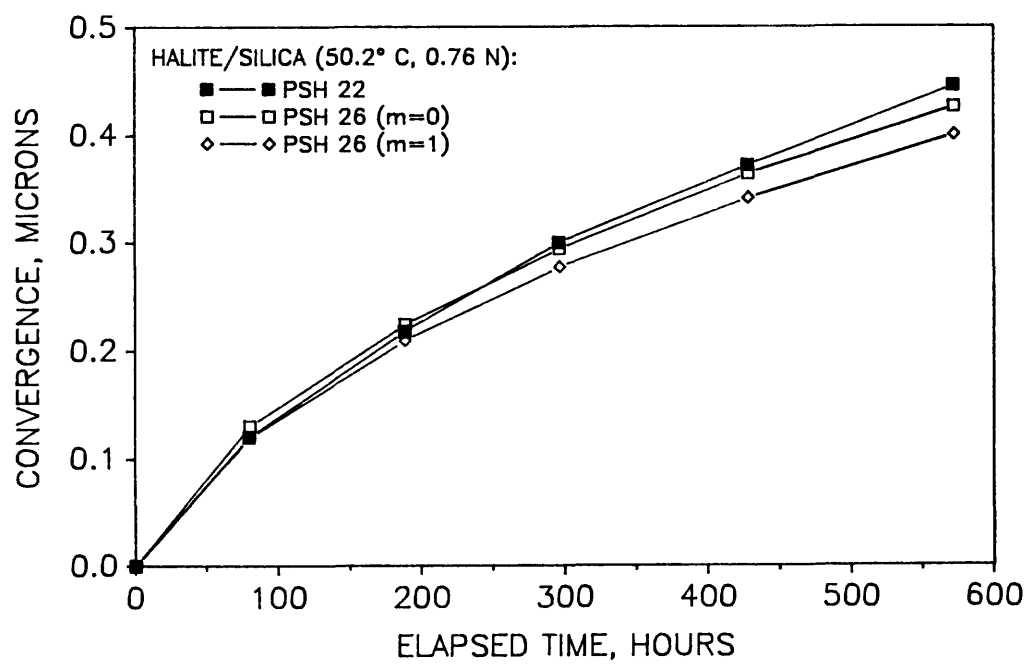



FIGURE 10: Comparison of convergence curves for two halite/silica experiments in saturated brine (PSH 22 and PSH 26) which were done under nearly identical conditions (see Table 1). Convergence in both experiments was determined using Equation 3, although only the zero-order minima (contact spot) and first-order minima were used in the analysis for PSH 26 (see text). Convergence curves from both of these orders are shown.

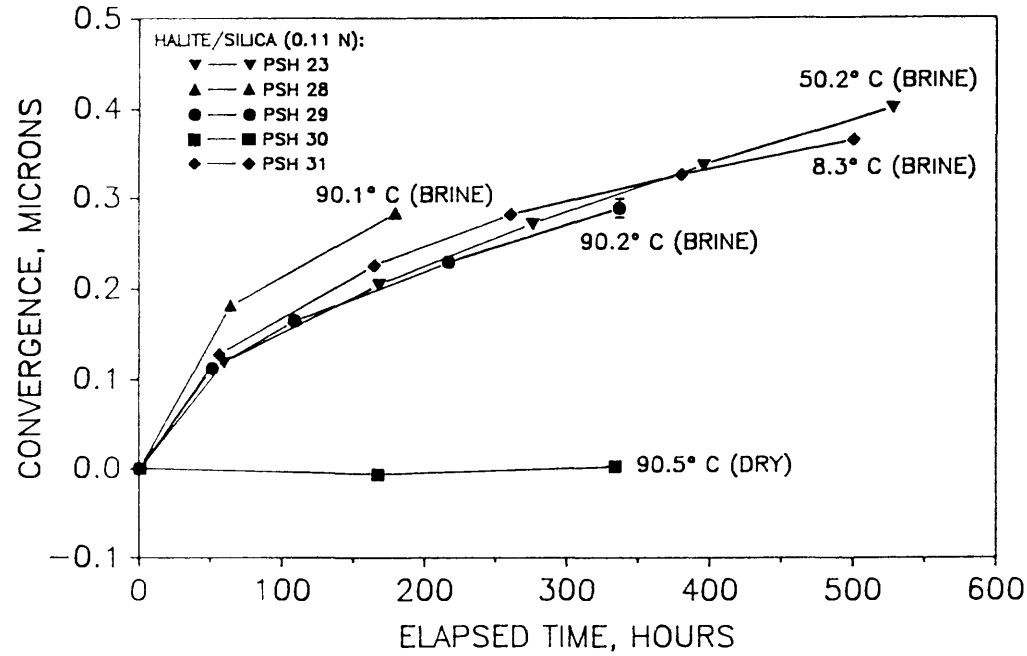



FIGURE 11: Convergence as a function of time for halite/silica experiments in saturated brine or dry nitrogen at 8.3°-90.5° C and loads of 0.11 N. The normal stresses and contact spot radii for all of these experiments are comparable (e.g. Figure 7B) and are not shown. With the exception of the final data point for PSH 29, the error bars (at ± one standard deviation) are approximately the same size as the symbols. In experiment PSH 28, convergence was measured using interference minima of m = 10 - 17 to avoid the damaged region near the contact spot (see Figure 6).

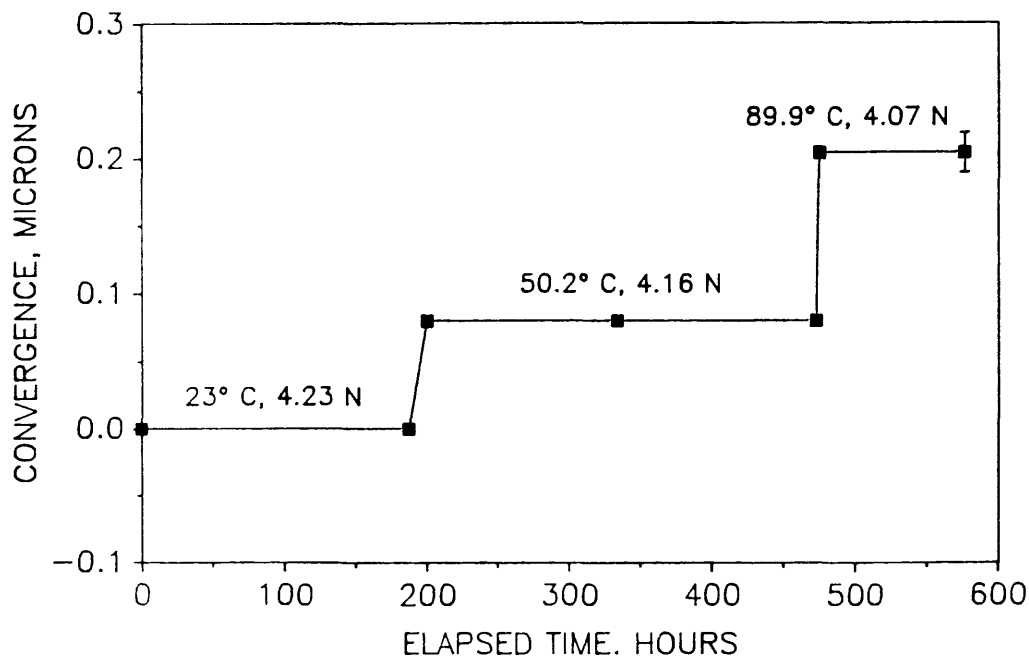



FIGURE 12: Convergence for a halite/silica experiment in moist air at 4.07-4.23 N (PSH 34), in which temperature was increased in a stepwise manner from 23°-89.9° C over 576 hours. The indicated decrease in load (corresponding to a decrease in $P_m$ from 13.4-11.2 MPa) is due to thermal expansion of the sample chamber and the temperature dependance of the spring calibration.

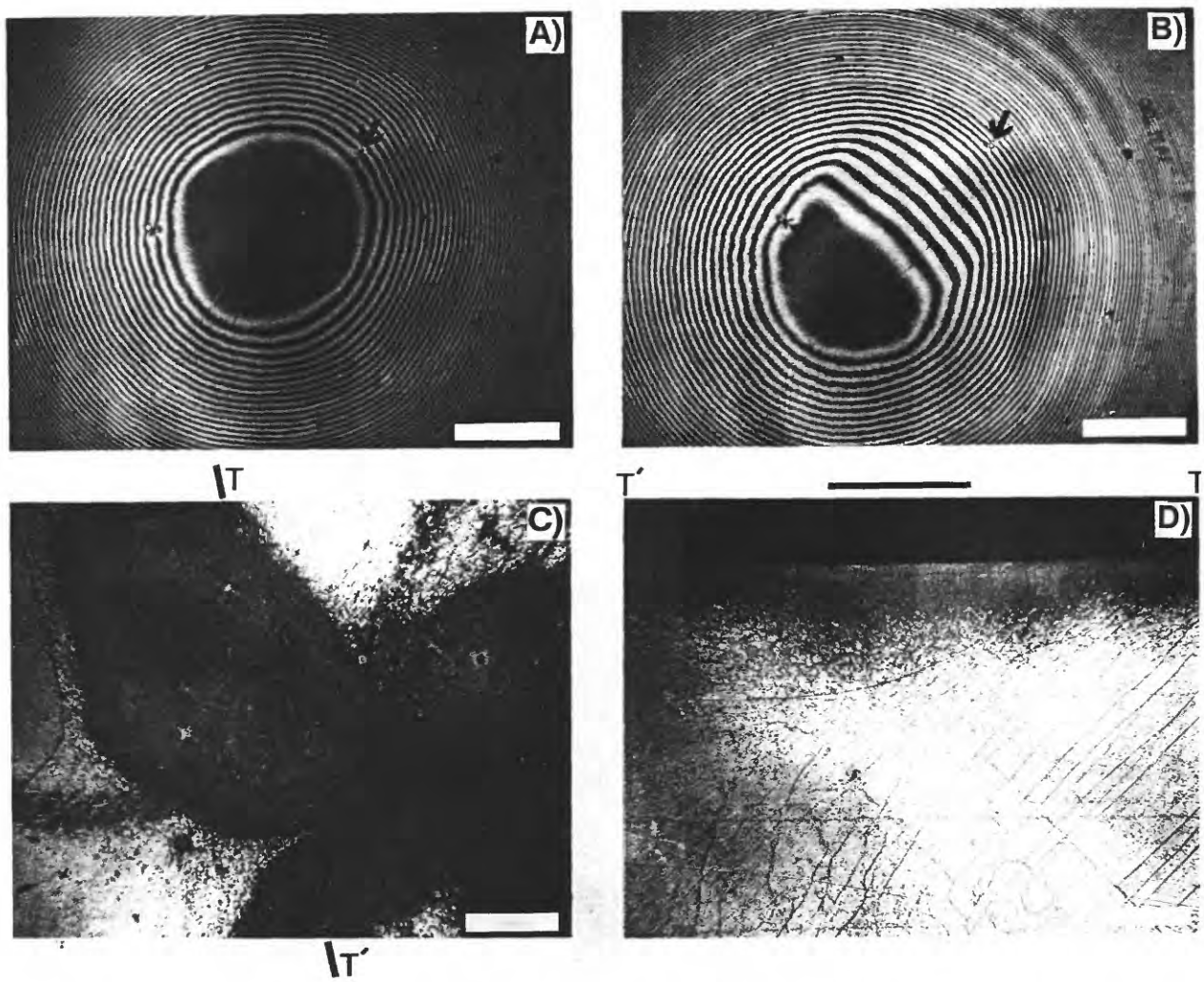



FIGURE 13: A) In situ sodium interferogram of sample PSH 34 at the conclusion of the experiment but before disturbing the microscope stage. B) Sodium interferogram of the same sample after partially opening the stage and tilting the window to reveal plastic flattening of the lens (shown by straight fringes in center of photo). As these photographs were taken in air, the change in lens separation between adjacent interference fringes is 0.295 μm. C) Reflected white light photomicrographs of PSH 34 after etching to reveal the emergence points of dislocations (dark spots). Dislocation rosettes produced by impressed grit particles are visible as small X-shaped patterns. One of these particles is highlighted by an arrow in these photographs to facilitate location of the contact spot. D) Cross section of PSH 34 after splitting along T-T' in C and etching. The size and position of the contact spot is indicated by the horizontal black bar. Black lines on the surface of the lens and within the interior of the crystal are sub-grain boundaries. The scale bars are 0.5 mm.

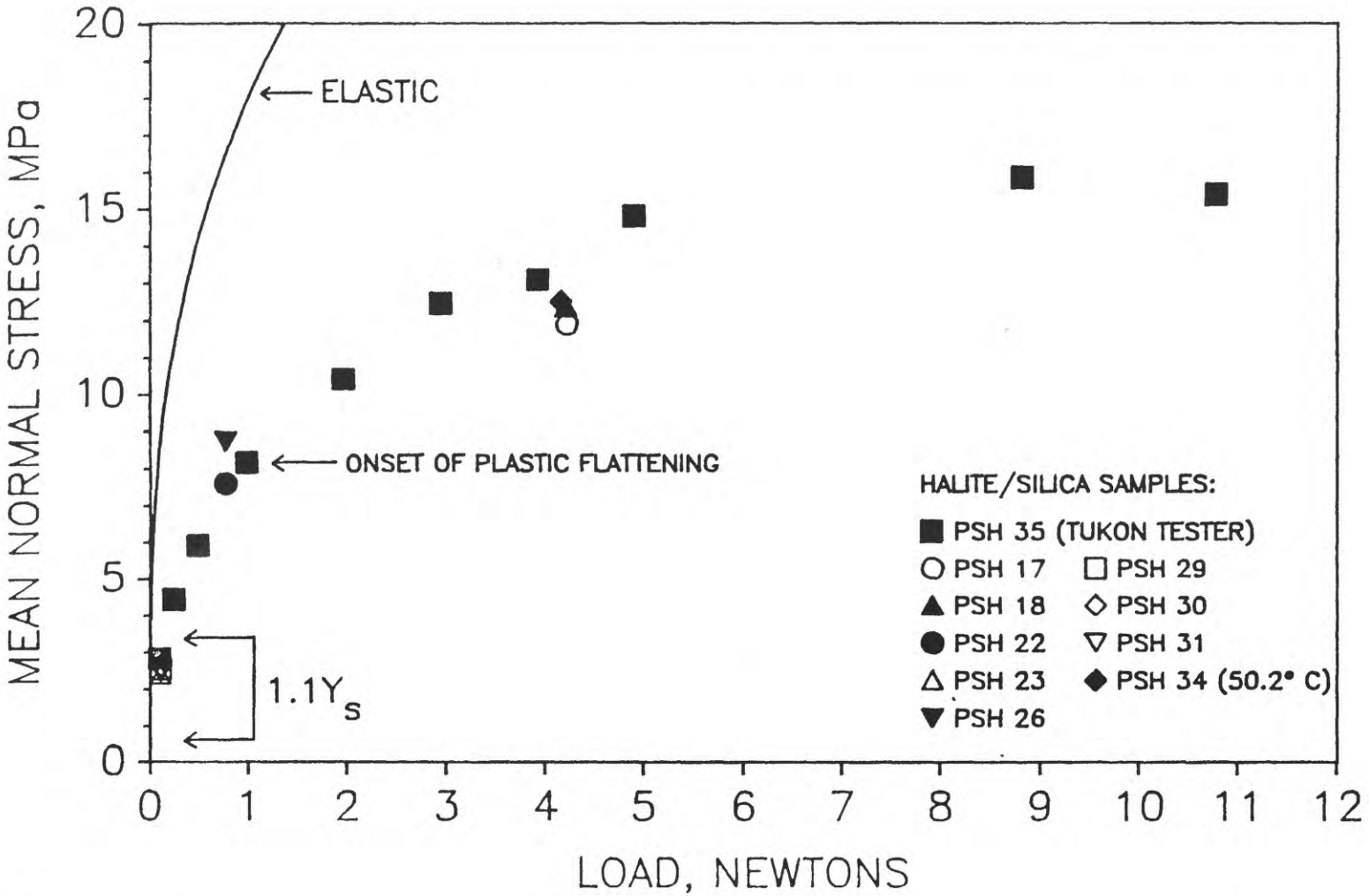



FIGURE 14: Mean effective normal stress, $P_m$, versus load from PSH 35 using the modified Tukon Hardness Tester shown in Figure 3. These measurements were made at 20° C and 50% relative humidity. The calculated variation of $P_m$ with load for the purely elastic contact of halite and silica lenses for the geometry used in this study is shown for comparison with the PSH 35 data. Also shown are $P_m$ from the halite/silica experiments; only $P_m$ at 50.2° C is plotted for PSH 34 to facilitate comparison with other experiments at the same load and temperature.

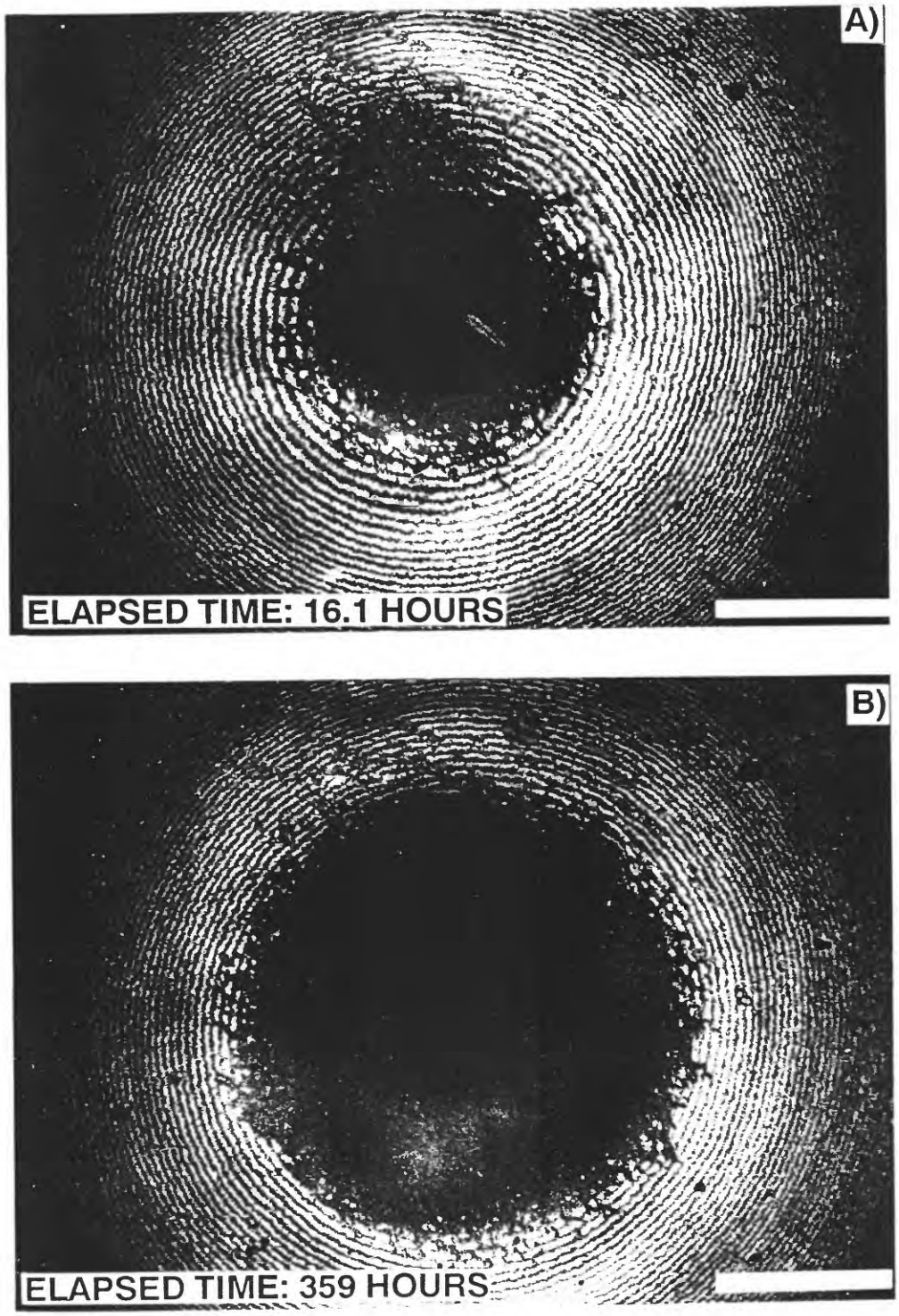



FIGURE 15: Reflected sodium interferograms of experiment PSH 27 in saturated brine at 50.2° C and 0.76 N for the times shown. In this experiment, a halite convex lens was pressed against a fused silica flat coated with an approximately 0.8-μm-thick layer of sodium montmorillonite to ascertain the effect of a clay interlayer on convergence rates. The change in lens separation between adjacent interference fringes is 0.214 μm. The scale bars are 0.5 mm.

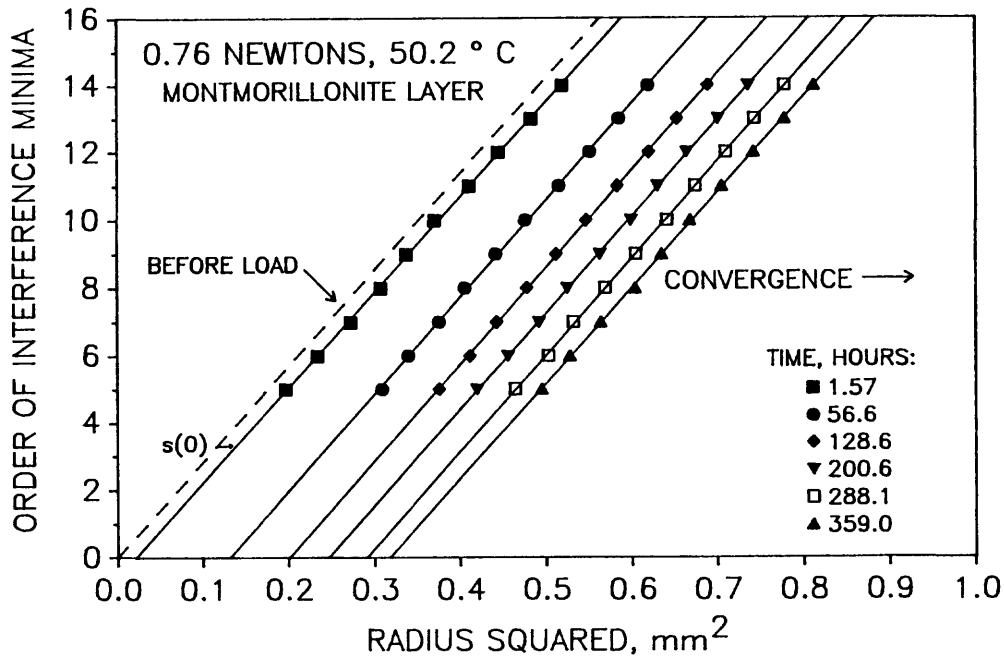



FIGURE 16: Plot used to measure convergence in the halite/clay/silica experiment, PSH 27, using Equation 2. The dashed line indicates the expected relationship between the interference minima order number, m, and the square of the radius of that minima, $r_m^2$, before the load is applied. The solid lines are based on linear regressions to all of the order numbers shown. The convergence that occurs immediately after the load is applied is indicated by s(0), and is followed by time-dependent convergence.

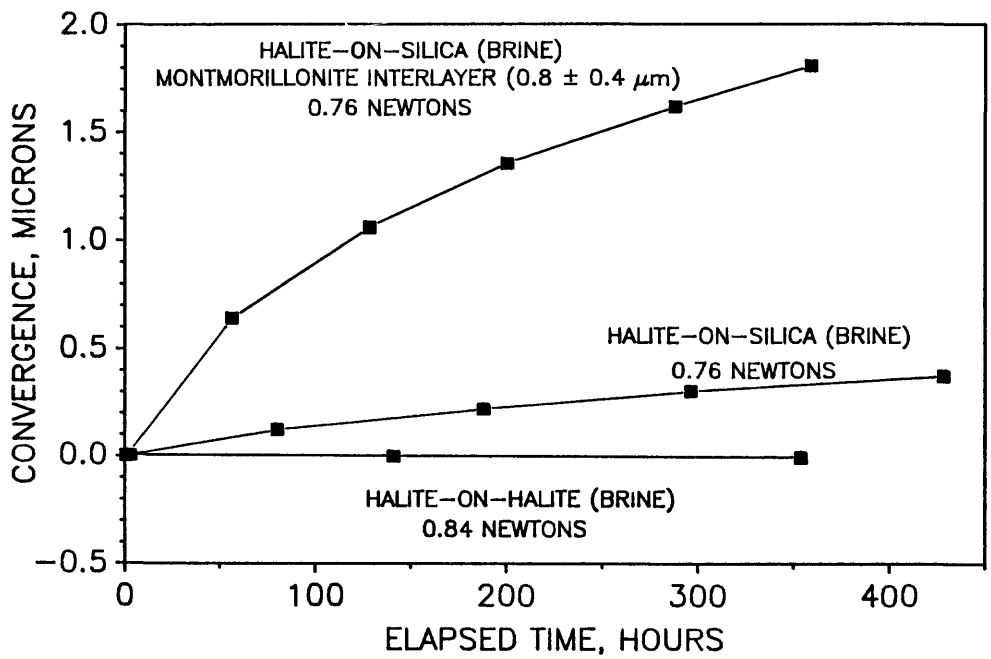



FIGURE 17: Convergence as a function of time for the halite/clay/silica (PSH 27) and halite/silica (PSH 22) experiments in saturated brine at 0.76 N and 50.2° C. A halite/halite experiment at 50.1° C and 0.84 N is shown for comparison (PSH 15; from Hickman and Evans, 1991). Notice the compressed vertical scale relative to Figure 9. The error bars (at ± one standard deviation) are approximately the same size as the symbols and are not shown.

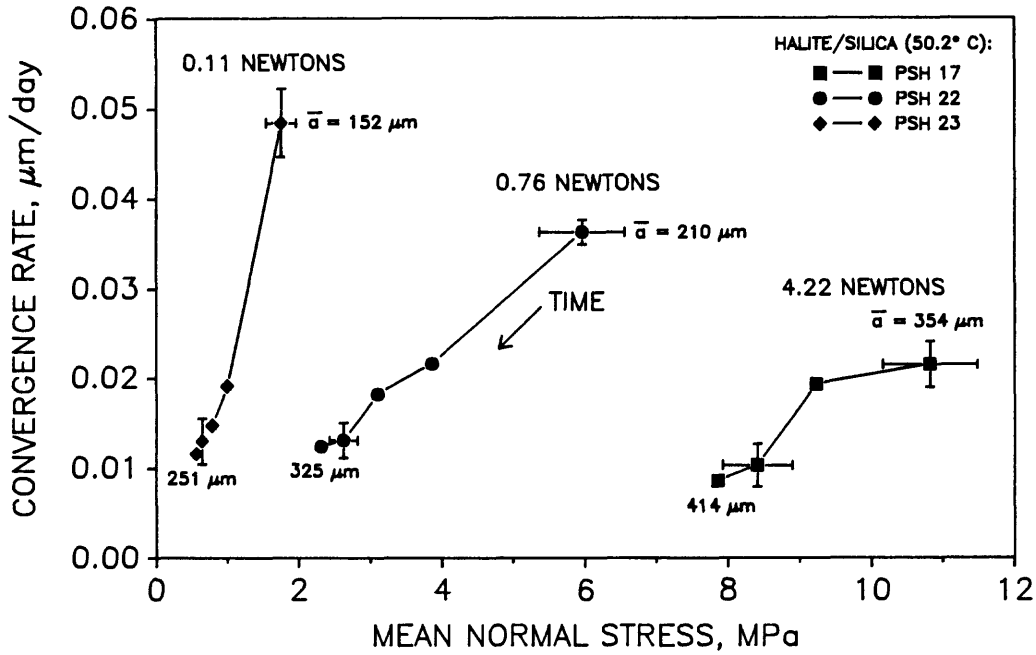



FIGURE 18: Instantaneous convergence rate as a function of $\overline{P_m}$ for the halite/silica experiments in brine shown in Figure 9. The data from a single experiment are connected by straight lines, with time increasing to the left. The mean contact spot radii, $\bar{a}$, during the initial and final time intervals for each experiment are shown. Representative uncertainties are given as the estimated measurement accuracy for $\overline{P_m}$ and plus-or-minus two standard deviations in convergence rate.

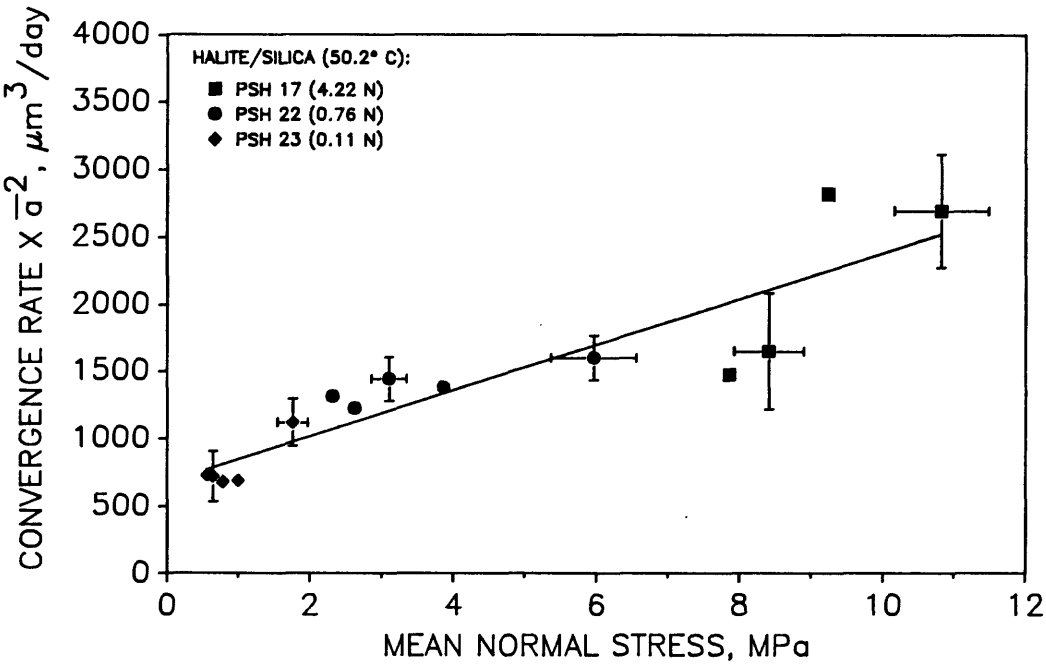



FIGURE 19: Convergence rates from Figure 18 multiplied by the square of the mean contact spot radius, $\bar{a}$, during each time interval as a function of $\overline{P_m}$. Representative uncertainties were calculated from the estimated measurement accuracy for $\overline{P_m}$ and $\bar{a}$ and plus-or-minus two standard deviations in convergence rate. The straight line is based on a linear regression to all the data.

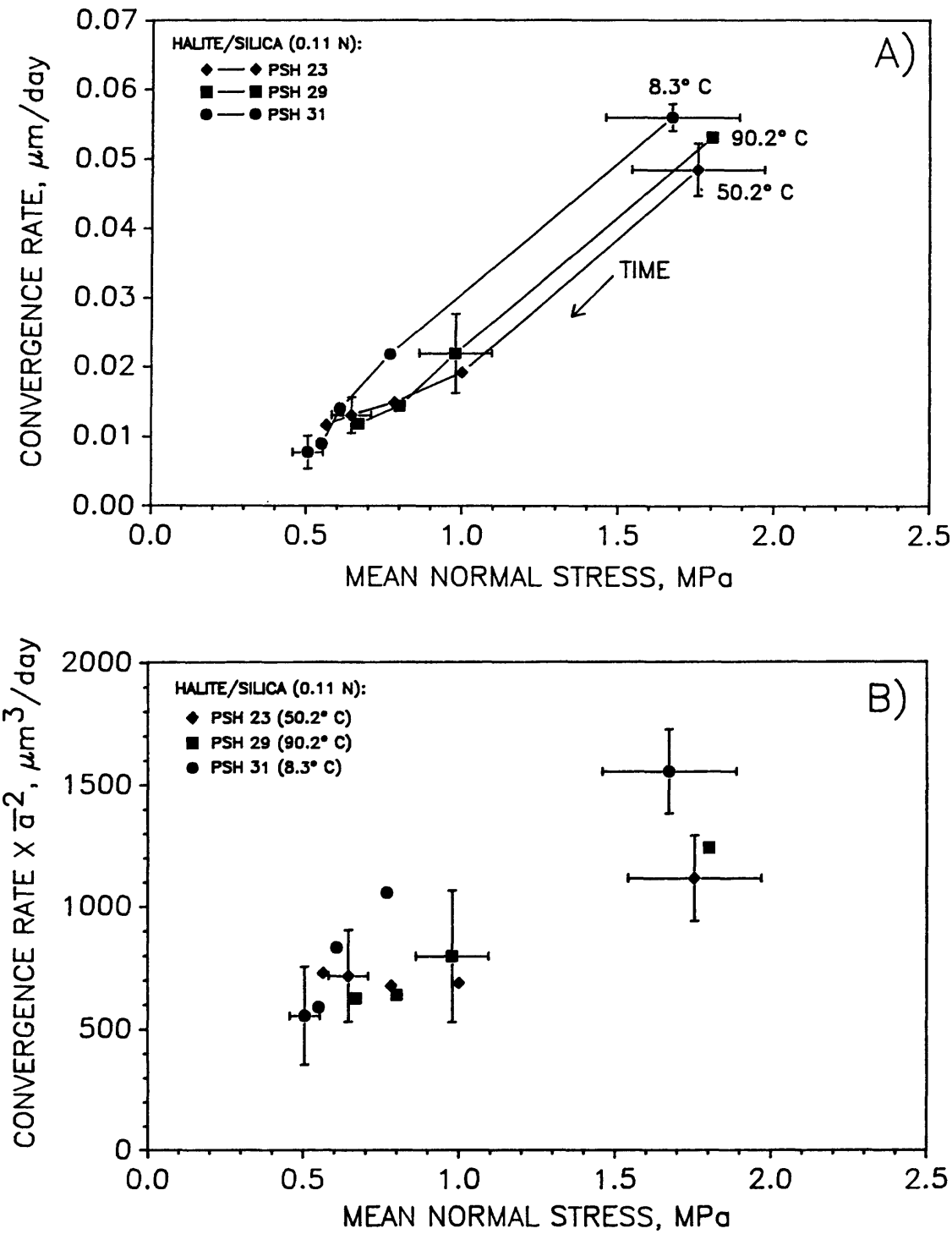



FIGURE 20: A) Instantaneous convergence rate as a function of $\overline{P_m}$ for the halite/silica experiments in brine at 0.11 N and 8.3°-90.2° C (see Figure 11). The data from a single experiment are connected by straight lines, with time increasing to the left. The mean contact spot radii are comparable for all three experiments and are not shown. Representative uncertainties were computed as for Figure 18. B) Convergence rates from A) multiplied by the square of the mean contact spot radius, a, during each time interval. Representative uncertainties were computed as for Figure 19.

# Rate laws for Water-Assisted Densification and Stress-Induced Water-Rock Interaction in Sandstones

Thomas Dewers
School of Geology and Geophysics, University of Oklahoma, Norman, OK 73019
and
Andrew Hajash
Department of Geology, Texas A&M University, College Station, TX 77843

**Abstract**

Mineral-water interactions under conditions of non-hydrostatic stress are thought to play a role in subjects as diverse as time-dependent creep in fault zones, phase relations in metamorphic rocks, mass redistribution in marl-limestone and sand-shale couplets, diagenetic replacement reactions, and the loss of porosity in deep sedimentary basins. As a step toward understanding the fundamental geochemical processes involved, using naturally rounded St. Peter sand, we have investigated the kinetics of pore-volume loss and quartz-water reactions under non-hydrostatic, hydrothermal conditions in flow-through reactors. Rate laws for creep and mineral-water reaction are derived from the time-rate of change of pore volume, sand-water dissolution kinetics, and (flow-rate independent) steady-state silica concentrations, and reveal functional dependencies on rates made by the variables of grain size, volume strain, temperature, effective pressure (confining minus pore pressure), and specific (quartz grain) surface areas. Together the mechanical and chemical rate laws form a self-consistent model for the coupled deformation/water-rock interaction of porous sands under non-hydrostatic conditions. Microstructural evidence shows a progressive widening of nominally circular and nominally flat grain-grain contacts with increasing strain or, equivalently, porosity loss. Taken together, the mechanical and chemical data suggest that the dominant creep mechanism in our experiments has been due to the removal of mass from grain contacts (termed pressure solution or solution-transfer). The magnitude of a stress-dependent concentration increase is too large to be accounted for by elastic- or dislocation-strain energy induced supersaturations invoked by many authors to be responsible for solution-transfer creep. In addition, number densities of observed ultra-fine particles and microfractures either present initially or produced at grain contacts during creep appear too low to account for the observed $c_{ss}$ increase. The most likely thermodynamic diving force that can account for the experimental results is the pressure dependence of the molar Gibbs free energy associated with grain interfaces. As current researchers favor a fluid-pressure dependence (and not normal stress) for this term, our results call for a reexamination of the thermodynamics of nonhydrostatically-stressed systems as are now in use.

## Introduction

One of the most important, and least understood, diagenetic processes leading to porosity modification in the subsurface is considered to be some form of solution-transfer creep (Rutter, 1983; Tada and Siever, 1989; Houseknecht, 1988; Spiers and Schutjens, 1990). It has been implicated, often circumstantially, in subsurface porosity loss (Houseknecht, 1984; 1988), as a creep process within fault zones (Rutter and Mainprice, 1978; Chester and Higgs, 1992), in the development of stylolites (Heald, 1955), and in providing solutes for cements (McBride, 1989). The precise operational dynamic governing solution-transfer creep is a subject of debate. There are two fundamentally different mechanisms currently in favor, which date back to classic studies by Weyl (1959) and Bathurst (1958). The Weyl hypothesis, which is generally equated to the process of "pressure solution", involves grain-boundary or thin-film diffusion in response to chemical potential gradients induced by normal tractions across grain-grain contacts. The

Bathurst model, also known as free-face pressure solution (Tada et al, 1987), strain solution (Wintsch and Dunning, 1985), or free-face dissolution (Engelder, 1982), involves reaction at pore-fluid stressed solid-fluid interfaces, or "free" faces, due to variations in strain energy and/or surface tension (Reuschle et al. 1988). In this model, the "free-faces" may exist at contact peripheries, within grain contacts in an "island-channel" structure, or both.

Variations on these themes have been proposed because of the several possible thermodynamic driving forces and kinetic mechanisms that act in concert to produce the reaction and transport. For example, "free-face" models combining essentially elastic strain energy with free-fluid diffusional transport have been proposed (Lehner, 1990; Spiers and Schutjens, 1990), as well as models involving stress-induced grain boundary diffusion within semicoherent grain interfaces (Gratz, 1991), or models comprising stress-induced dissolution kinetically limited by interface reaction (Raj, 1982). Authors have recognized that surface-energy controlled processes act in concert, either competitively with or enhancing stress- or strain energy-controlled processes (Raj, 1982; Pharr and Ashby, 1983). These would include grain-boundary wetting (Watson and Brennan, 1987) and crack healing (Brantley et al, 1990; Hickman and Evans, 1991). At the center of the "pressure solution" controversy is the question of existence of continuous adsorbed fluid films at grain boundaries. Tada et al. (1987) discuss disjoining pressure of thin films between quartz plates using data by Pashely and Kitchener (1979), but these films are not stable (at least those in equilibrium with an external atmospheric pressure) above temperatures of 70°C (Derjaguin, 1971).

In addition, intense microfracturing of grains beneath rigid indenters or at grain contacts increase solubility (Tada et al., 1987), aided possibly by the formation of glass along associated surfaces of frictional sliding (Ferguson et al., 1987). The phenomena of stress-corrosion cracking (Atkinson, 1984) is often linked to pressure solution, but the precise thermodynamics and kinetics associated with this process are not clear.

In this paper we attempt to provide an empirical basis for the model most applicable to solution-transfer creep in clean quartzose sand. Results of sand compaction experiments are discussed, in which fluid chemistry and reaction rates have been monitored simultaneously with deformation. As such, these experiments differ from conventional ones in which fluid-rock interaction is examined post-deformation, as in studies of the influence of crystal lattice defects on reaction kinetics (Brantley et al., 1986; Blum et al., 1990, MacInnis and Brantley, 1992). We examine influences of applied stress on dissolution rates and quartz "solubility" (we use the term solubility rather loosely here, to describe the flow-rate independent steady-state concentration of aqueous silica; at conditions of zero effective stress, this becomes the concentration of aqueous silica in equilibrium with quartz at conditions of fluid pressure and temperature). Based on our results, we argue for the operation of a stress-dependent, rather than strain-dependent, thermodynamic driving force involved in diagenetic solution-transfer during at least the initial stages of lithification. This contrasts with recent theoretical work (Bruton and Helgeson, 1983; Lehner, 1990; Dahlen, 1992). Using principles of mass conservation, we propose a simple dual-site reaction kinetics model to account for the observed chemical changes induced by the applied stress. This model suggests that an observed increase in solubility with increasing effective pressure ($P_e$; defined as the difference between confining and pore pressures) is due to a trade-off between the different thermodynamics and reaction kinetics acting at the different mineral sites (i.e. loaded faces and "free" faces). We find that the magnitude of stress-influenced quartz solubility increases with decreasing grain size, for constant $P_e$, temperature, and volume strain. Because reaction kinetics are proportional to surface area or surface density of active sites (which change with grain size), this particular dependence underscores the likelihood that the concentration behavior is influenced by kinetic factors and does not reflect an "equilibrium" situation. This type of behavior is also consistent with grain size-silica importer/exporter relationships described by Tada and Siever (1989), Porter and James (1986), and Houseknecht

(1988), in which finer grained sands appear to act as exporters of mass to coarser grain sands. Diffusional transport down dissolved silica gradients observed in our experiments could produce mass transfer from coarser to finer sands.

An in-depth treatment and summary of the experimental data and modeling attempts described in this study is given in Dewers and Hajash (in preparation). Here our purpose is to concentrate on the precise mechanisms active during compaction by giving an overview of our results and modeling attempts.

## Driving Forces for Solution-Transfer Creep

Tada et al. (1987) summarize the various thermodynamic driving forces thought to attend each mechanism for solution-transfer, expressed in terms of the activity of silica in quartz (using a standard state of unit activity at conditions of fluid pressure and temperature). Table 1 shows some of their results at conditions of 300 °K and 30.0 MPa effective pressure. The first three mechanisms listed can be grouped together as "free-face" processes, and the fourth can be termed a "loaded-face" process. Notice that the dislocation strain- and elastic strain-induced changes as have been studied by Engelder (1982), Wintsch and Dunning (1985), and others, yield a very small increase, while that due to microgranulation (and here we refer to the production of ultra-fine particles with high surface area to volume ratios, and thus high and non-negligible contributions to overall free energy of reaction made by surface energy), and the "loaded-face" mechanism yield higher contributions under conditions of interest here. Although the driving force for "loaded-face" dependent solution-transfer is two to three orders of magnitude greater than the "free-face" type, Tada et al. (1987) have suggested that the kinetics of the free-face process yields a faster overall rate of deformation. Coupled to the thermodynamic driving forces listed in Table 1, the kinetics of solution-transfer processes may be controlled by grain boundary diffusion (Gratz, 1991), diffusion through hydrostatically stressed pore fluid (Lehner, 1990; Lehner and Bataille, 1985; Tada et al., 1987; Spiers and Schutjens, 1990), diffusion through adsorbed fluid films having substantially different properties than a bulk fluid phase (Rutter, 1983; Rutter and Mainprice, 1978; Hickman and Evans, 1991), or interfacial reaction (detachment or attachment rates) (Raj, 1982; Mullis, 1991).

Elias and Hajash (1992) examined the influence of effective pressure on steady-state (flow-rate independent) silica concentrations in distilled water reacting with St. Peter sand. Under an effective pressure of 34.5 MPa, the silica concentration increases roughly 10% above that at zero effective pressure. If it is fair to compare silica concentrations in pore fluid to theoretical predictions of silica activities in quartz (strictly speaking, this is a crude estimate), it is obvious that only the latter two mechanisms listed in Table 1, ultra-fine particle production through grain crushing and loaded facet-related reaction, could in principle produce an effect of a magnitude similar to that observed in the Elias and Hajash experiments. We investigate this observed solubility increase further in this paper.

## Experimental Design and Procedures

The materials used and the experimental design surmount some shortcomings of other experimental systems utilized in wet compaction experiments. We use naturally rounded sands, which avoid the angular nature of crushed and sieved quartz crystals commonly used (e.g. Schutjens, 1991). Confining pressure is applied isotropically, not directionally as is done by most experimentalists attempting to reproduce "pressure solution" in the laboratory (Schutjens, 1991; Hicks and Applin, 1988; Gratier and Guiget, 1986). Because of both of these factors, we avoid much of the microcracking commonly observed in low temperature compaction experiments.

*Starting Materials* - Unconsolidated St. Peter sand (Ordovician Cave Unit, Minnesota) was used in all experiments reported in this study (as well as in the experiments of Burns, 1989, and Elias and Hajash, 1992). The sand was mechanically sieved and the .09-.12 mm (very-fine), .18-

.25 mm (fine), and .25-.35 mm (medium) fractions were retained for starting material. In order to remove iron oxide contaminants, the sand was subjected to an acid wash in a 25% HCl solution for one hour, then rinsed repeatedly in deionized distilled water. This was followed by rinsing (up to 6 hours) in an ultrasonic bath, to faciliate removal of any ultra-fine particles adhering to grain surfaces. Specific surface areas were then determined by the single point B.E.T. method, using krypton as the absorbate gas. These are reported in Table 2. The pore fluid used in all experiments was deionized distilled water.

*Hydrothermal Flow-Through System*- The flow-through system utilized in this study has been described in detail in Hajash and Bloom (1991). Five such systems were used during the course of this study. The system centers around a stainless steel pressure vessel, with inlet lines for confining fluid and thermocouples, and inlet and outlet ports for pore fluid lines. Within the vessel, the sample is separated physically from the confining fluid by flexible cylindrical jackets, composed of gold or Teflon™. This allows an effective stress to be established across the sample. Both PFA and TFE type heat-shrink Teflon™ roll covers were used as jackets in this study. The TFE jackets were found to degrade appreciably at temperatures near 200°C and so were avoided at these conditions. After jacketing, the sample geometry is one of a 4.4 cm diameter cylinder 8 to 9 cm in length. The pore fluid pressure and flow rates are controlled by an ISCO syringe pump which can maintain either constant flow or constant pressure. Parts of the system in contact with pore fluid at temperature are constructed of corrosion resistant alloys (either Hastelloy™or Inconel™), or titanium. Confining pressure is maintained by an air-driven pump. The vessel is placed within a three zone furnace; internal temperature of the reaction cell is measured with two K-type thermocouples. Once targeted (uniform) pressure and temperature conditions were achieved, the jacketed sample was subjected to several "pore-fluid flushes." A pore-fluid flush consists of passing several pore volumes (≈ 150 ml) of distilled water through the sample over a short period of time (≈ 2 hour) so that the "reacted" pore fluid is removed. These are used to ensure dissolution of "fines" and in aging sand surfaces prior to the beginning of sampling, and to lower the concentration of silica in the pore fluid in order to study reaction kinetics. In many cases, the sample is exposed to temperature and fluid pressure conditions over the course of several weeks, with zero applied effective pressure, to promote aging of the quartz grain surfaces prior to the application of effective pressure.

*Sampling Procedures* - The high solid-surface area to pore-fluid volume ratio and large sample volume of the experimental design allow the determination of both solubility and kinetic parameters during the course of a single experiment simply by varying flow rate. Lower limits to flow rates were set by the sampling frequency and sample volume. A typical case would involve removing four to five mls of pore fluid (which are replaced simultaneously at constant pressure with the ISCO pump) one to three times per week. Higher, near-constant flow rates (up to 1 ml per minute) were used to monitor dissolution kinetics. During sampling, a small volume (~2 ml) was removed to "purge" the pore lines, followed by removal of up to 6 additional mls. Pore fluid was extracted using a hand-operated pressure generator with a small reservoir coil; constant pressure was maintained by the ISCO pump during fluid withdrawal. Portions of sample were acidified with .5 ml 25% HCl solution, diluted with deionized distilled water, and retained for analysis in Nalgene™ bottles. Frequently, pH of the effluent solution at ambient room temperature was measured.

*Analytical Techniques* - Silicon concentrations in pore fluid samples were determined using atomic absorption spectrophotometry with matrix-corrected standards. Samples from one experiment at 200°C were analyzed for fluorine (from the Teflon™ jackets) using ion chromatography, and were found to contain less than 5 ppm. Similar experiments using the same apparatus have yielded total iron and titanium concentrations below detection limits (Burns, 1989).

## Mechanical Results and Model

*Determination of Pore Volume Loss* Amounts of pore volume loss are determinable during the course of an experiment by measuring the amount of fluid extracted from the pressure vessel to maintain constant pressure conditions. This value is corrected for density changes and excess volumes of the back-pressure system used during the extraction. We use definitions for strain ($\varepsilon$) and strain rate ($\dot{\varepsilon}$) that follow common usage (e.g. Schutjens, 1991):

$$\text{Volume Strain} = \varepsilon = \frac{\phi - \phi_o}{1 - \phi} = \frac{dV}{V_o}$$

$$\text{Volumetric Strain Rate} = \dot{\varepsilon} = \frac{1}{V}\frac{dV}{dt}$$

where $\phi$ is porosity, $\phi_0$ is initial porosity, dV is an increment of (pore) volume change (pore volume decrease is taken to be positive), $V_0$ is initial bulk volume, and t is time.

An example of raw volume strain versus time data is shown in Figure 1. This experiment utilized very-fine grained sand, a temperature of 150 °C, fluid pressures of 34.5 MPa, and confining pressures that vary from 34.5 to 69.0 MPa. The sample was subject to three phases of loading, between which the effective pressure was reduced to zero. Compaction commences for effective pressures greater than zero; it ceases when effective pressure is removed. Much of the initial volume strain associated with loading the sample to a desired effective pressure and recoverable pore volume during unloading is due to elastic compression of the sand-water mixture, and behaves similarly to that found in porous sandstone compressibility experiments (Zimmerman et al., 1986). Scrutiny of the time-dependent loss of porosity in Figure 1 reveals that the rate of strain decreases through time, or with accumulating volume strain. Because of the hydrostatic loading configuration of our experimental apparatus, and because grain volumes are roughly conserved during the experiments, we will use the terms volume strain, total strain, and pore volume strain interchangeably.

All volume strain-time data sets were fit to the following expression using a commercially available software package; this may not be the only mathematical expression that will provide a good "fit" to some of the data, but it is the only one we have found that represents data from all of the experiments to an excellent degree of correlation:

$$\varepsilon - \varepsilon_0 = a * \ln\{b + (t - t_0)\} - a * \ln\{b\} \qquad (1)$$

where $\varepsilon_0$ is the volume strain at time $t_0$, and 'a' and 'b' are coefficients determined by the least squares regression method. An example comparing observed versus predicted volume strain using (1) of reduced strain ($\varepsilon - \varepsilon_0$) versus time ($t - t_0$), is shown in Figure 2. Note that the model fits the observed data very well. Figure 3 shows volume strain behavior for an experiment using fine-grained sand at 200°C, pore fluid pressures of 34.5 MPa, and effective pressures from 0 to 48.3 MPa. The curves in this figure are model extrapolations back to zero strain, using (1), which show what the strain-time pathway would have been if the effective pressure would have been maintained constant through time. The rate of strain, or slope of the tangents to the curves, increase as the strain approaches zero. We define the slope of the curve at zero volume strain as the "limiting strain rate". While this is a fictional strain rate (i.e. there is always some elastic or other type of strain associated with achieving an effective pressure of appreciable magnitude), we find the limiting strain rate to be a useful quantity in comparing results from many different experiments, because it factors out the influence of strain on strain rate. This is discussed further below.

Because the overall strain rate decreases with increasing strain, the creep in our experiments can be said to exhibit "strain hardening". The manner in which the sand-water system undergoes strain hardening is probably indicative of the actual creep mechanism. In Figure 4, we show results from an experiment examining halite creep in water performed by Hickman and Evans (1991). They pressed polished lenses of halite against flat glass plates under controlled conditions and examined the manner in which the halite deformed. They found, for the case shown in Figure 4, that the halite lens approached the glass plate, or converged, by removal of mass at the grain contact and associated contact flattening. The solid curve in Figure 4 is a fit to Hickman and Evan's (1991) data using a convergence-time relation analogous to (1). The change in convergence (denoted 's') at a single grain contact may be related to overall volume strain in an aggregate consisting of similar grains of dimension L by the relation:

$$\varepsilon_v = \frac{n(s - s_0)}{L_0}, \qquad (2)$$

where n is the total number of grain contacts, $L_0$ is initial grain size, and $s_0$ is the initial distance between centers of adjacent grains. It follows by comparing (1) and (2) that the strain hardening parameters a' and a are related by

$$a' = a\frac{L_0}{n}. \qquad (3)$$

Here we can begin to gain a physical significance for the curve fit parameter 'a' in (1). We suggest that the strain-hardening observed experimentally in our aggregate experiments is consistent with the single contact behavior observed in the Hickman and Evans (1991) experiments, and is ultimately due to growth of contact area, although there are some differences inherent in each experimental design that may hinder direct comparison (e.g. we use constant stress, or effective pressure, in our experiments, while Hickman and Evans use constant load).

*Constitutive law for initial sand compaction* Differentiating the expression in (1) above with respect to time, one finds (for $t_0 = 0$)

$$\frac{d\varepsilon}{dt} = \frac{a}{b+t} \quad \text{or} \quad = \frac{a}{b} * \exp\left\{-\frac{(\varepsilon - \varepsilon_0)}{a}\right\} \qquad (4)$$

when using (1) to substitute for t. This shows that the rate of volume strain decreases with the exponential of volume strain. This is a useful expression, because we have found that the product

$$\frac{a}{b}\exp\left\{\frac{\varepsilon_0}{a}\right\}$$

is essentially constant in experiments utilizing similar grain size, temperature and effective stress, and in experiments which have experienced several loading and unloading episodes. Here $\varepsilon_0$ is understood to be both elastic and permanent volume strain at the beginning of a loading period at constant effective pressure. This means that we can compare results between experiments at different total volume strains, because (4) allows one to separate the time rate of change of total strain into strain-dependent and strain-independent parts. Letting $\dot{\varepsilon}_{lim}$ represent the "limiting" rate of volume change, i.e. that rate at zero strain (we call this the limiting strain rate), (4) can be expressed as

$$\frac{d\varepsilon}{dt} = \dot{\varepsilon}_{lim}\exp\left(\frac{-\varepsilon}{a}\right) \qquad \text{where } \dot{\varepsilon}_{lim} \equiv \frac{a}{b}\exp\left(\frac{\varepsilon_0}{a}\right). \qquad (5)$$

We see from (5) that the limiting strain rate is just the rate of volume strain accumulation at zero strain, as we discussed above in reference to Figure 3.

With the ultimate goal being the derivation of a constitutive law for the observed sand compaction, we have explored the influence of effective pressure, temperature, and grain size on the limiting strain rate (these being variables identified as being important in chemical compaction processes, see for example Houseknecht, 1984). In Figure 5 we plot the limiting strain rate from a suite of experiments (using fine grained sand) as a function of effective pressure at 150° and 200°C. An exponential relationship fits the data in both plots to an excellent degree. Figure 6 expresses the temperature dependence of the limiting strain rate from three experiments run at constant grain size (fine) and effective pressure (34.5 MPa). This is a typical Arrhenius plot; a linear relationship in the space of log(rate) versus reciprocal temperature suggests an exponential dependence of rate on temperature. An apparent activation energy of 79 kj/mol is indicated by a least-squares fit to the data. Finally, in Figure 7, we plot the effect of grain size on the limiting strain rate; a linear increase is evident with inverse grain size.

In theoretical treatments of diffusion creep or pressure solution, it is common to associate a linear dependence between rates of volume strain and effective pressure (i.e. a linear viscous relationship). Our data would suggest, on the other hand, a linear dependence only at low stress, with a nonlinear dependence becoming evident at higher stresses. A similar relationship was derived theoretically by Dewers and Ortoleva (1990a, b). The inverse temperature dependence with the logarithm of limiting strain rate is common in thermally activated processes, although we caution that other functional dependencies on temperature may fit our data just as well (see Brantley et al., 1989 for a discussion of this). The inverse grain size dependence is consistent with theoretical models as well, and with the grain size effect observed by Houseknecht (1984; this was shown by Dewers and Ortoleva, 1990b). We will return to the physico-chemical significance of how each variable influences the strain rate after we discuss the chemical results.

The above empirically derived relationships have led us to propose the following constitutive law between volumetric strain rate $\dot{\varepsilon}$, effective pressure, grain size, and temperature:

$$\dot{\varepsilon} = \frac{A}{L_o(1-\varepsilon)} \exp\left(-\frac{E_a}{RT}\right)\left\{\exp(\alpha P_e) - 1\right\}\exp(-h\varepsilon) \qquad (6)$$

where

$A$ = pre-exponential factor, equal to ~ 2 cm-hr$^{-1}$
$E_a$ = apparent activation energy, equal to 79 kj/mol
$L_O$ = initial grain size (in cm)
$R$ = gas constant
$\alpha$ = parameter measuring the strength of the effective pressure dependence, and
$h$ = "strain-hardening coefficient", dependent on temperature, grain size, and effective pressure; equal to 1/a.

The (1-$\varepsilon$) factor in the denominator of (6) results from the differences between $\frac{d\varepsilon}{dt}$ and $\dot{\varepsilon}$ (i.e. the inverse dependence on V or $V_O$ for $\dot{\varepsilon}$ and $\frac{d\varepsilon}{dt}$ respectively), and physically accounts for the decrease in grain dimension at grain contacts accompanying compaction by removal of mass. This type of constitutive law is a necessary ingredient in theoretical studies of time dependent compaction (e.g. Angevine and Turcotte, 1983; Dewers and Ortoleva, 1990 a;b; Mullis, 1991), however, we must emphasize the preliminary nature of this rate law, as other forms may fit the data equally as well. For example one would expect from theoretical treatments that the $\alpha$ or 'h'

parameters may vary inversely with temperature. Such claims are probably not yet justifiable given our present data set, and should be explored in future experiments.

## Chemical Results and Model

*Enhanced "solubility" and reaction kinetics with applied effective stress.* Elias and Hajash (1989; 1992) found an enhanced solubility that varied proportionately and semi-reversibly with effective pressure. In the present study, we have attempted to reproduce their results, and, in analogy with the above treatment of compaction creep, investigate further any effects of grain size, temperature, and volume strain. We focus on these variables because they are often examined in creep experiments to discern activation energies and kinetic mechanisms for solution-transfer or other non-dislocation associated mechanisms for creep (Raj, 1982; Cooper and Kohlstedt, 1984; Spiers and Schutjens, 1990; Schutjens, 1991). We discuss two of these experiments in detail, leaving the remainder to Dewers and Hajash (in prep.), and propose a dynamic model to account for the observations.

Figure 8 summarizes the results of a year-long experiment, also shown in Figure 1, utilizing very-fine grained sand. Here an effective pressure of 34.5 MPa was applied over the course of a week, maintained for approximately 500 hours, subsequently removed, and then reapplied; this process was then repeated. Several fluid flushes were performed, as described above, to monitor dissolution rates throughout the cycle. As before, the silica concentration (which, at the rate of fluid extraction used, roughly 5-10 mls per week, did not vary with flow rate) increases directly with effective pressure. However, upon removal of effective pressure, followed by a pore-fluid flush, the flow-rate independent silica concentration at zero effective pressure was the same as that prior to loading. Upon reapplication of effective pressure, the concentration increased to a level slightly higher than during the first application. Then, following another fluid flush, the concentration in the pore fluid obtained a level quite close to the value obtained during the first period of loading. This suggests that the sand solubility at zero and 34.5 MPa effective pressure obtains a characteristic and reproducible value. After the second removal of $P_e$, silica concentration was slightly higher than the pre-loading value. This slowly decreased through time, approaching quartz saturation at temperature and fluid pressure conditions (160 ppm). Thus the concentration increase appears to be reversible, especially at low strains.

A noticeable decrease in steady-state silica concentration with time is apparent with constant $P_e$, especially during the third loading episode in Figure 8. One possible explanation for this is the increase in volume strain that occurs simultaneously with the observed concentration change. It is the only system variable that changes during the loading phases; an increase in volume strain is commensurate with increasing contact area, so that the concentration increase could be reflecting a decrease in magnitude of normal stress acting at grain-grain contacts.

A run at 200°C using fine grained sand is depicted in Figure 9, showing the concentration behavior with time and effective pressure, which was increased step-wise during the experiment. During the long phase at constant effective pressure of 20.7 MPa (between ~1400 and 2300 hours) the concentration decreases slightly, again probably due to increasing volume strain, but increases each time the effective pressure is increased, at nearly constant volume strain.

Influence of $P_e$ and grain size on steady-state silica concentration Summaries of the influences of effective pressure and grain size on the (flow rate-independent) steady-state aqueous silica concentration at low strains can be seen in Figures 10 and 11. Figure 10 shows that the % increase in silica concentration above quartz saturation increases exponentially with increasing effective pressure; at 200°C, the curve begins to flatten with increasing volume strain (compare the solid and long-dashed curves in Fig. 10) at low effective pressure. A simple relation involving the exponential of effective pressure has been used to fit these data; the rationale behind this equation will be discussed further below. Figure 11 shows how the steady-state silica concentration changes with grain size at constant

effective pressure and temperature and nearly constant strain. It is our belief that this grain size dependence originates from kinetic factors. Consider that a porous aggregate consists of both "loaded" and "free" grain interfaces, and that both types of interface may be sites for either dissolution or precipitation reactions. As grain size decreases, the available surface area *per grain* for precipitation decreases, so although dissolution rate at grain contacts may or may not change with grain size (see below), the precipitation rate will decrease with decreasing grain size. If the bulk pore fluid concentration equals the ratio of forward to backward reaction rates (as is the case for equilibrium between quartz and water at zero-effective pressure conditions), decreasing the precipitation (backward reaction) rate with all else held constant should lead to an increase in the bulk level of aqueous silica concentration. A model accounting for these effects is presented below. The observed grain size dependence of silica concentration, if also existing within pores of a porous sand body buried at depth, could have important implications for import and export of silica between sandstones of different grain sizes during burial diagenesis.

Reaction kinetics as a function of $P_e$, grain size, and $\varepsilon_v$ If the steady-state silica concentration rises with effective pressure, it is then natural to expect that there might be an effective pressure dependence on the dissolution kinetics of sand in water as well. Figure 12 shows raw concentration-time data from experiments using (a) very-fine, (b) fine, and (c) medium grained sand. The two data sets in each plot refer to the concentration-time behavior at 0.0 and 34.5 MPa effective pressure (both at low volume strains). The data for $P_e$ = 0.0 MPa (filled circles) have been fit to a first-order kinetics solution from a model involving quartz sand reaction and silica transport along a reaction column with ideal plug flow, which is outlined in the next section. The rate of silica concentration increase with time is roughly proportional to the rate of sand dissolution. It is seen in all three case that rates are faster (i.e. greater initial slope in concentration versus time) under conditions of effective pressure compared to rates at $P_e$ = 0.0 MPa. Figure 13 attempts to quantify this "excess" rate by plotting a normalized rate of dissolution (if we denote the rate under conditions of effective pressure as $R^+_{P_e}$, and that at zero effective pressure as $R^+_0$, then the normalized rate we use here is defined as $\{R^+_{P_e} - R^+_0\} / R^+_0$) as a function of effective pressure. As was the case for the % increase in [$SiO_2$] and in creep rate, a simple equation involving the exponential of effective pressure fits the data set for the fine grained sand at 150°C very well. It is interesting to note that although we found a grain size dependence to the % increase in steady-state silica concentration, the extent to which stress-enhanced dissolution rate is influenced by grain size is less obvious. It appears that a single curve might fit all 150°C data in Figure 13 equally as well, including very-fine, fine, and medium sands. This underscores our belief that the grain-size dependence of the steady-state silica concentration arises from variations in precipitation kinetics due to grain size and not from the $P_e$-influenced dissolution rate. We believe that the enhanced dissolution rates reflect dissolution from grain contacts, while the background rate reflects dissolution from grain free faces. This latter value should not change appreciably with effective pressure (as long as total volume strains, and thus surface area changes, are small).

## A Dual-Site First-Order Kinetics Model

In this section, we set forth a quantitative model for effective pressure-dependent quartz sand-water interactions. This model could be used in conjunction with the constitutive law given by (6) to obtain a self-consistent model for quartz sand-water interaction and deformation under burial conditions, and to predict rates of porosity loss and/or cementation accompanying various hydrologic and structural regimes. Direct connections of this model with thermodynamic, kinetic, and geometric factors are discussed further in Dewers and Hajash (in prep.). We need only briefly discuss such connections here.

*Reaction at conditions of zero effective pressure* A general equation accounting for changes in aqueous silica concentration at a point within a reaction column due to quartz-water reaction and advective mass transport is given by

$$\frac{\partial(c_\alpha \phi)}{\partial t} = -\vec{\nabla} \cdot (c_\alpha \vec{q}) + \nu_\alpha R_\alpha . \qquad (7)$$

where $c_\alpha$ = molar concentration of dissolved silica (or dominant silicon species existing in solution), $\phi$ is porosity, q is volume flux, $\nu_\alpha$ is the stoichiometric number relating moles of silica released per mole of quartz dissolved, and $R_\alpha$ is rate of overall quartz sand-water reaction. Under the assumptions of perfect plug flow in a cylindrical reactor, the volume flux, volumetric flow rate, and macroscopic fluid velocity parallel to the long axis of the reaction column are related by

$$\frac{QL}{V_b \phi} = \frac{q}{\phi} = v \qquad (8)$$

where
Q = volumetric flow rate through the reactor (volume/time)
$L$ = length of cylindrical reactor
$V_b$ = bulk volume of reactor (solid plus fluid)
v = linear fluid velocity (length/time).

The rate of reaction, $R_\alpha$, is found from considerations of reaction kinetics in the system. At conditions of hydrostatic pressure, for example, a rate law valid for quartz-water reaction at a single type of quartz-water interface is expressed by (Rimstidt and Barnes, 1980; Dove and Crerar, 1990):

$$R_{SiO_2\text{-}qtz} = n\rho A k^+ a_{SiO_2} (a_{H_2O})^2 \left(\frac{Q}{K} - 1\right)$$

where
n = number density of quartz grains per unit volume
$\rho$ = solid molar density of quartz ( = .044 mol/cm$^3$)
A = average surface area per grain
$k^+$ = dissolution rate constant (length per time)
$a_{SiO2}$ = activity of quartz (defined to be unity at standard state conditions of pure quartz at hydrostatic pressure and temperature),
$a_{H2O}$ = activity of water (unity at infinite dilution), and
Q/K = saturation state of the system where K is the equilibrium constant for quartz-water reaction at the standard state conditions defined above, and

$$Q = \frac{a_{H_4SiO_4}}{(a_{SiO_2})(a_{H_2O})^2}$$

Note that the product (nA) divided by the porosity is the total solid surface area to fluid volume ratio. The rate equation is written in this form for convenience in describing water-rock reaction in granular media. (nA) can be found from BET-determined specific surface areas and knowledge of the solid mass density of quartz. Considering that, for precipitation rate constant $k^-$, equal to

$$k^- = \frac{k^+}{K},$$

we may re-write the above rate law as:

$$\frac{R}{\phi} = R^- c - R^+ \qquad (9)$$

where $R^-c$ is the precipitation rate and $R^+$ the dissolution rate:

$$R^- = \frac{n\rho}{\phi} Ak^-$$

$$R^+ = \frac{n\rho}{\phi} Ak^- c^{eq}$$

Here, $c^{eq}$, equal to $R^+/R^-$, is the concentration of aqueous silica in equilibrium with quartz under the conditions of interest (we have assumed unit activity of water and unit activity coefficient for aqueous silica, which is valid for quartz reacting in distilled water under hydrostatic pressure conditions).

With this (7) becomes for constant flow rate:

$$\frac{\partial c}{\partial t} = -\frac{QL}{V_b \phi}\frac{\partial c}{\partial z} - R^- c + R^+ - c\frac{\partial \ln(\phi)}{\partial t} \qquad (10)$$

In (10), z is the Cartesian coordinate parallel to the long axis of the reaction column. This equation is easily solvable if the flow rate, sample volume, porosity, $R^-$ and $R^+$ may be considered constant. These are valid assumptions over short time scales in which porosity and surface areas do not change appreciably.

We have used a flow-rate stepping procedure to determine quartz-water reaction kinetics in our experimental system. Basically this involves changing the rate of fluid flow along the column, and allowing the concentration to come to steady-state. The steady-state silica concentration is characteristic of the flow rate at fast flow rates; for sufficiently long times and large residence times (equivalent to low flow rates), the concentration at the column outlet (we denote this as $c(L,t)$ where $L$ is the column length) approaches $\frac{R^+}{R^-} = c^{eq}$. If we let the fluid flow velocity characteristic of the system at steady-state as 'u', and 'v' as the flow velocity of a new state, then the concentration will respond to the change in flow rate from u to v as (for times less than one residence time) as (Dewers and Hajash, in prep.)

$$c(L,t) = c^{eq} - (c^{eq} - c_0)\exp\{-R^-(1-\frac{v}{u})t\}. \qquad (11)$$

Rearranging terms in (11), one finds

$$\ln\left(\frac{c - c^{eq}}{c_0 - c^{eq}}\right) = -R^- t(1-\frac{v}{u}). \qquad (12)$$

If $v \ll u$, (12) reduces to expressions used previously by Rimstidt and Barnes (1980) and Murphy and Helgeson (1989) for determining $R^-$. An expression similar to (11) valid for times greater than the residence time at the new flow rate is found in Posey-Dowty et al. (1986), and is a steady-state at the new flow rate. $R^-$ can be found from raw concentration-time data, such as were given in Figure 12, using (11), and $R^+$ is then found by multiplying $R^-$ by $c^{eq}$.

*Rates of dissolution under conditions of effective pressure* Casting the kinetics of quartz-water reaction discussed above in terms of the rate of change of grain volume, one finds:

$$\frac{\partial V_g}{\partial t} = Ak^{-}(c - c^{eq}) \tag{13}$$

If we now account for two types of reaction sites on the grain surface, f (for free-faces) and c (for contacts), and assume that reaction mechanisms at both site types reflect first-order kinetics, (13) now becomes:

$$\frac{\partial V_g}{\partial t} = A_f k_f^{-}(c - \tilde{c}_f) + A_c k_c(c - \tilde{c}_c) \, . \tag{14}$$

where $\tilde{c}$ refers to a limiting concentration in the kinetic expression which may or may not correspond to an "equilibrium" concentration for a specified quartz-water interface, and is characteristic of a given mechanism for solution transfer by reaction at grain contacts. Dewers and Ortoleva (1990a;b) have shown that both interface-detachment controlled and diffusion controlled rates for quartz dissolution at grain contacts may be expressible in terms of first-order rates, if the appropriate k and $\tilde{c}_c$ are used. In this case, $R^+$ and $R^-$ used in (10), corresponding to dissolution and precipitation rates respectively, now become

$$\begin{aligned} R^- &= \frac{n\rho}{\phi}\{A_f k_f + A_c k_c\} \\ R^+ &= \frac{n\rho}{\phi}\{A_f k_f \tilde{c}_f + A_c k_c \tilde{c}_c\} \end{aligned} \tag{15}$$

The steady-state concentration in the pore-fluid, which we may tentatively term the "solubility" of quartz grains under applied effective stress, may still be found from (10) for sufficiently low flow rates (long residence times) and amounts of volume strain to be:

$$c_{ss} = \frac{R^+}{R^-} = \frac{\{A_f k_f \tilde{c}_f + A_c k_c \tilde{c}_c\}}{\{A_f k_f + A_c k_c\}} \tag{16}$$

Notice that, in contrast to conditions of zero effective stress wherein the steady-state concentration at long residence times approaches the equilibrium value, under conditions of effective stress the steady-state concentration at sufficiently long residence times now depends on kinetic, thermodynamic, and geometric parameters. As a consequence, the steady-state concentration becomes dependent on grain size (or, more likely, specific surface area). By defining a parameter, $\beta$, equal to:

$$\beta = \frac{A_f k_f}{A_c k_c} \tag{17}$$

(16) becomes

$$c_{ss} = \frac{\beta \tilde{c}_f + \tilde{c}_c}{\beta + 1} \tag{18}$$

If we assume a normal stress dependence for the increase in chemical potential above the standard state of hydrostatic pressure, $\tilde{c}_c$ may be expressed in the following form (this is derived in Dewers and Ortoleva, 1990a; b; see also Gratz, 1991):

$$\tilde{c}_c = c^{eq} \exp\left\{\frac{P_e f}{\rho RT}\right\} \qquad (19)$$

where $P_e$ is the macroscopically applied (or far-field) effective stress, f is a stress magnification factor which accounts for stress-focusing at grain contacts, R is the gas constant, and T is temperature. Combining (18) and (19) yields, if $c^{eq} \approx \tilde{c}_f$,

$$\frac{c_{ss}}{\tilde{c}_f} - 1 = \frac{\exp\{\frac{P_e f}{\rho RT}\} - 1}{\beta + 1} \qquad (20)$$

This is a useful form because it represents the deviation in steady-state silica concentration above quartz saturation due to applied effective stress, as was found in the Elias and Hajash (1992) experiments, and in our experiments. The expression in (20), although derived theoretically using some assumptions that have appeared elsewhere in the literature, is a similar equation to that used to fit the data in Figure 10. One can now see how the coefficients calculated in the least-squares fits of concentration data (e.g. in Figures 10 and 13) have a precise thermodynamic or kinetic meaning. The $\alpha$ parameter used in curve fitting in these two figures may in fact be interpreted as

$$\alpha = \frac{f}{\rho RT} \qquad (21)$$

and the $\beta$ parameter given by (17) can be interpreted as the ratio of reaction rates at free faces and grain contacts. Using coefficients calculated for 150°C and 34.5 MPa effective pressure, a typical value for f is calculated to be ~7, meaning that the local normal effective stress is roughly 7 times greater than the applied effective pressure. This is considerably less than the local stresses obtained from Hertzian elastic contact theory (Zhang et al., 1990), which give 'f' factors of 50 or greater depending on porosity. Thus solution-transfer creep can, at even low amounts of volume strain, reduce magnitudes of local stresses. Note that both f and $\beta$ are likely dependent on volume strain as well, however these dependencies may be of secondary importance based on our initial results.

A more detailed discussion of the thermodynamics and kinetics of water-rock interaction under conditions of nonhydrostatic effective pressure is given by Dewers and Hajash (in prep.), but we should point out some important considerations. The implications of (19) are far reaching; if indeed the local thermodynamic driving force (and, it follows, the Gibbs free energy along a grain contact or activity of silica in quartz at stressed interfaces) for solution transfer is influenced by a normal stress, then the thermodynamic models of Bruton and Helgeson (1983), Lehner and Bataille (1985), and Dahlen (1992) are fundamentally flawed and need to be reworked. It would seem from the data that both dissolution and precipitation rates influence water-rock interaction. If so, then the theoretical assertions of authors who examine solution-transfer problems in light of a single rate-controlling mechanism, such as Mullis (1991), need to be reexamined as well.

## Observations of Experimentally Produced Microstructure

A series of photomicrographs depicting fine grain sand solid run products after three experiments are shown in Figure 14. These are from thin sections prepared by injecting epoxy into the sand packs upon removal from the pressure vessel at the end of each particular experiment. In order to disturb the sand pack as little as possible, the epoxy was injected while the sand pack was still wet; although some original grain textures are likely preserved,

unfortunately this procedure led to less-than perfect results, and undoubtedly grains were moved around relative to one another. Nevertheless, Figure 14 (a) shows how a fine grained sand pack looks after going through a single loading event to 34.5 MPa effective pressure. There are some visible fluid inclusion trails, and the grains appear to "float" within the epoxy matrix. Figure 14 (b) is a fine grained sand pack subject to 34.5 MPa effective pressure and 200°C for ~3000 hours. Note that the overall pore volume space has been reduced, and that several flat grain-grain contacts are apparent. There are many fluid inclusion trails (suggesting healed microfractures), probably present at a higher density than in (a), but every grain does not contain such a trail. Figure 14 (c) is a sand sample that was subjected briefly to effective pressures above 55 MPa, which at the temperature of this experiment (200°C) was sufficient to produce grain crushing at grain contacts, similar to the "Hertzian" fracturing discussed by Zhang et al. (1990). Thus sands subject to long periods of effective pressure below the critical level for crushing undergo a time-dependent volume strain that consists at least in part in grain contact flattening, but may have a component associated with microcrack growth and healing.

With regard to changes in solubility accompanying grain crushing, we sampled pore fluid solution from the experiment depicted in Figure 14 (c) shortly after the crushing event, and found that it had increased more than 50% above quartz saturation at these conditions. After about 100 hours following the fracturing, the concentration of silica in the pore fluid had decreased to a level coinciding with quartz saturation. Thus while the ultra-fine particles produced by microfracturing can lead to very large increases in solubility (Tada et al., 1987), their effect is not long lasting, and certainly could not account for the sustained concentration increases that were evident in our experiments. While certainly many ultra-fine shards still exist in the sample depicted in Figure 14 (c), they are not present in sufficient amounts (or number density) to influence the pore fluid chemistry.

## Reconciling Mechanical and Chemical Results

A general relationship between compaction rate and dissolution rate was set forth by Dewers and Ortoleva (1990b), and was discussed by Gratz (1991). The volumetric strain rate of an aggregate of uniform grain size and volume strain undergoing isotropic compaction by a dissolution-dominated mechanism can be related to the dissolution rate at grain contacts by (Dewers and Ortoleva, 1990b)

$$\dot{\varepsilon}_v = \frac{3G_c}{L} \tag{22}$$

where $\dot{\varepsilon}_v$ is the volumetric strain rate, $G_C$ is the contact dissolution velocity, and L is grain size. (22) exhibits a dependence on inverse grain size that is consistent with that seen in Figure 7. Any further grain size dependence in solution-transfer creep laws derives from any grain size dependence of the dissolution velocity. In the case of diffusion-control, the dissolution velocity is inversely proportional to the cross-sectional area for diffusion, while there is no further dependence in the case of reaction control. In an earlier section, we related molar dissolution rates under local nonhydrostatic and hydrostatic conditions to general grain shape changes by the following equations:

$$\begin{aligned} R^+_{EP} &= \frac{n\rho}{\phi}\left(A_c G_c + A_f G_f\right) \\ R^+_0 &= \frac{n\rho}{\phi}\left(A_f G_f\right) \end{aligned} \tag{23}$$

where $R^+$ is dissolution rate in mol/area-time (subscripts "EP' and '0' refer to conditions of effective pressure, and zero effective pressure, or alternatively nonhydrostatic and hydrostatic conditions, respectively), n is the number of grains per unit volume, $\rho$ is quartz molar density, $\phi$

is porosity, A is surface area per grain, and G is reaction velocity at the subscripted interfacial type ('c' denotes contacts, 'f' denotes free faces). To a first approximation in which we neglect any change in free-face area during the initial stages of compaction, the dissolution velocity at grain contacts may be found by subtracting these two rates and using (23)

$$G_c = \left(R_{EP}^{+} - R_0^{+}\right)\frac{\phi}{n\rho A_c}. \qquad (24)$$

Finally, by expressing the grain number density in terms of the grain volume and porosity ($n = (1-\phi)/V_g$) and substituting (24) into (22), we can express volumetric strain rates in terms of experimentally determined reaction rates:

$$\dot{\varepsilon}_v \cong \frac{3\phi V_g}{(1-\phi)\rho A_c L}\left(R_{EP}^{+} - R_0^{+}\right) \qquad (25)$$

Table 3 compares volumetric strain rates calculated from rates of pore volume loss to those calculated from measured dissolution kinetics and equation (25). It is significant that strain rates from "mechanical data" all lie within a factor of four of estimates made from "chemical data", suggesting that the observed volume strain could have been produced by dissolution at grain contacts. This points out the strength of the experimental design, in that both volume loss and reaction rates can be measured simultaneously; in the case of isotropic compaction by solution-transfer, each data set is an independent manifestation of the same physico-chemical phenomena.

It is worth noting that the lack of any grain size dependence measured for $G_C$ is consistent with the inverse grain size (and not the cube of grain size) dependence of the volumetric strain rates, by the form of equation (6) above. This would suggest either that dissolution velocities for grain shortening by solution-transfer, $G_C$, are controlled by reaction kinetics (see rate laws used by Dewers and Ortoleva, 1990b) or that the cross sectional area for diffusion does not change significantly with increasing grain size (Gratz, 1991).

## Summary and Conclusions

The main conclusions of our study are:

- St. Peter sand subject to conditions of effective pressure at elevated temperatures and in the presence of water undergoes time-dependent compaction in a regular and predictable manner. We have proposed a "creep" law accounting for our experimental data, in which volumetric strain rates increase proportionately with the exponential of effective pressure and temperature, and inversely with grain size. Creep rates decrease systematically with increasing volume strain, thus the system appears to "strain-harden". This creep law, although still somewhat preliminary, can be used in models of porosity evolution during burial diagenesis.

- The concentration of silica dissolved in pores of sandstones is a function of both lithostatic (rock) and hydrostatic (fluid) pressures, and increases as an exponential function of the effective pressure (difference between lithostatic and hydrostatic pressures). This likely reflects a trade-off between the differing kinetics and thermodynamic driving forces acting at sand grain facets loaded hydrostatically and those transmitting lithostatic load. We conjecture that although some microfracturing is observed in our experiments, the volume density of produced fines or new surface area is too low to induce much change in the bulk concentration of silica in the pore fluid. A heightened solubility has been observed in experiments in which quartz grains were

intentionally crushed, but the effect disappears rapidly compared to the time scale of sustained concentration increases observed in our experiments.

- The magnitude of the percent increase due to effective pressure increases slightly with decreasing grain size. Although small, when applied over geologic time, this increase would result in mass transfer (down concentration gradients) from finer grained sands to more coarse grained sands. The added cement to coarse-grained sands would inhibit their further compaction, while the mass lost from finer sands, because it does not cement pores, may have an enhancing effect on compaction rates. This differential cementation and compaction in sands of heterogeneous grain size has been discussed by Porter and James (1986); Houseknecht (1988) and Tada and Siever (1989).

- Dissolution kinetics of quartz sand in distilled water are enhanced by the application of diagenetically reasonable effective stresses. The magnitude of the enhancement increases with the exponential of effective pressure (similar to the solubility and volume strain rate described above). This increase is presumed to be due to an additional mechanism for dissolution acting at non-hydrostatically stressed grain contacts, which likely differs in thermodynamics and kinetics from the mechanism responsible for dissolution in distilled water at hydrostatically stressed quartz-water interfaces. This additional mechanism, because it is stress-related, constitutes a process for "pressure solution".

## Acknowledgments

The experimental portion of this research was performed while the senior author was a post-doctoral fellow at the Center for Tectonophysics, Texas A&M University. We need thank Brian Elias for valuable help in the lab and John Logan for advice, such as on the merits of PFA teflon. Some of the experiments were performed on a hydrothermal system on loan from AMOCO Production Co. in Tulsa, OK. Support from the National Science Foundation (EAR-9219709), a contract from the Gas Research Institute (to John Logan), and start-up funds from the University of Oklahoma are gratefully acknowledged.

## References

Angevine, C., L. and Turcotte, D.L., 1983, Porosity reduction by pressure solution: a theoretical model for quartz arenites: Geological Society of America Bulletin, v. 94, p. 1129-1134.

Atkinson, B., 1984, Subcritical crack growth in geologic materials: Journal of Geophysical Research, v. 89, p. 4077-4114.

Bathurst, R. G. C., 1958, Diagenetic fabrics in some British Dinantian limestones: Liverpool Manchester Geological Journal, v. 2, p. 11-36.

Blum, A.E., Yund, R.A., and Lasaga, A.C., 1990, The effect of dislocation density on the dissolution rate of quartz: Geochimica et Cosmochimica Acta, v. 54, p. 283-297.

Brantley, S., Crane, S.R., Crerar, D.A., Hellmann, R., and Stallard, R., 1986, Dissolution at dislocation etch pits in quartz: Geochimica et Cosmochimica Acta, v. 50, p. 2349-2361.

Brantley, S., Evans, B., Hickman, S.H., and Crerar, D.A., 1990, Healing of microcracks in quartz: Implications for fluid flow: Geology, v. 18, p. 136-139.

Bruton, C.J., and Helgeson, H.C., 1983, Calculation of the chemical and thermodynamic consequences of differences between fluid and geostatic pressure in hydrothermal systems: American Journal of Science, vol. 283-A, p. 540-588.

Burns, G.L., 1989, The solubility of natural quartz sand at 100°C and 150°C, 345 bars: An experimental investigation in a flow-through system: M.S. Thesis, Texas A&M University.

Chester, F.M., and Higgs, N.G., 1992, Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions: Journal of Geophysical Research, v. 97, p. 1859-1870.

Cooper, R.F., and Kohlstedt, D.L., 1984, Solution-precipitation enhanced diffusional creep of partially molten olivine-basalt aggregates during hot-pressing: Tectonophysics, v. 107, p. 207-233.
Dahlen, F.A., 1992, Metamorphism of nonhydrostatically stressed rocks: American Journal of Science, Vol. 292, p. 184-198.
Dewers, T.A., and Hajash, A. (1991) Mechanical and chemical compaction of natural quartz sand: An experimental and theoretical study: Geological Society of America Abstracts with Programs, Vol. 23.
Dewers, T.A., and Ortoleva, P., 1990a, Geochemical self-organization III: a mean-field pressure solution model of metamorphic differentiation: American Journal of Science, v. 290, p.473-521.
Dewers, T. A., and Ortoleva, P., 1990b, A coupled reaction/transport/mechanical model for intergranular pressure solution, stylolites, and differential compaction and cementation in clean sandstones: Geochimica et Cosmochimica Acta, v. 54, p. 1609-1625.
Dove, P.M., and Crerar, D.A., 1990, Kinetics of quartz dissolution in electrolyte solutions using a hydrothermal mixed flow reactor: Geochimica et Cosmochimica Acta, v. 54, p. 955-969.
Elias, B. P., and Hajash, A., 1989, Compaction of unconsolidated St. Peter sand under non-zero effective stress: an experimental investigation: Geological Society of America Abstracts with Programs, v. 21., p. 10.
Elias, B.P., and Hajash, A., 1992, Changes in quartz solubility due to effective stress: An experimental investigation of pressure solution, in review, submitted to *Geology* on 3/91.
Engelder, T., 1982, A natural example of the simultaneous operation of free face dissolution and pressure solution: Geochimica et Cosmochimica Acta, v. 46, p. 69-74.
Ferguson, C.C., Lloyd, G.E., and Knipe, R.L., 1987, Fracture mechanics and deformation processes in natural quartz: a combined Vickers indentation, SEM, and TEM study: Canadian Journal of Earth Science, Vol. 24, p. 544-555.
Gratier, J.P., and Guiget, R., 1986, Experimental pressure solution-deposition on quartz grains: the crucial effect of the nature of the fluid: Journal of Structural Geology, p. 845-856.
Gratz, A.J., 1991, Solution-transfer compaction of quartzites: progress toward a rate law: Geology, v. 19, p. 910-904.
Hajash, A. and Bloom, M., 1991, Marine Diagenesis of Feldspathic Sand: A Flow-Through Experimental Study at 200°C, 1 kbar: Chemical Geology, v. 89, p. 359-377.
Heald, M.T., 1955, Stylolites in sandstones, Journal of Geology, v. 63, p. 101-114.
Hickman, S.H., and Evans, B., 1991, Experimental pressure solution in halite: the effect of grain/interphase boundary structure: Journal Geological Society London, v. 148, p. 549-560.
Hicks, B. and Applin, K., 1987, Experimental pressure solution of quartz sands at T< 100°C: preliminary results: AAPG Research Conference on Prediction of Reservoir Quality Through Chemical Modeling (I. Meshri and D.L. Gautier, conveners), Park City, Utah, June, 1987.
Houseknecht, D. W., 1984, Influence of grain size and temperature on intergranular pressure solution, quartz cementation, and porosity in a quartzose sandstone: Journal of Sedimentary Petrology, v. 54, p. 348-361.
Houseknecht, D.W., 1988, Intergranular pressure solution in four quartzose sandstones, Journal of Sedimentary Petrology, v. 58, p. 228-246.
Lehner, F.K., 1990, Thermodynamics of rock deformation by pressure solution, *in* (Barber, D.J. and Meredith, P.G., eds.) Deformation Processes in Minerals, Ceramics and Rocks, Unwin Hyman, London, p. 296-333.
Lehner, F.K., and Bataille, J., 1985, Nonequilibrium thermodynamics of pressure solution, Pure and Applied Geophysics, v. 122, p. 53-85.
MacInnis, I.N., and Brantley, S.L., 1992, The role of dislocations and surface morph-ology in calcite dissolution: Geochimica et Cosmochimica Acta, v. 56, p. 1113-1126.
McBride, E.F., 1989, Quartz cement in sandstones: A review: Earth-Science Reviews, Vol. 26, p. 69-112.

Mullis, A.M., 1991, The role of silica precipitation kinetics in determining the rate of quartz pressure solution: Journal Geophysical Research, v. 96, p. 10007-10013.
Murphy, W.M., and Helgeson, H.C., 1989, Thermodynamic and kinetic constraints on reaction rates among minerals and aqueous solutions. IV. Retrieval of rate constants and activation parameters for the hydrolysis of pyroxene, wollastonite, olivine, andalusite, quartz, and nepheline: American Journal of Science, Vol. 289, p. 17-101.
Pashley, R.M. and Kitchener, J.A., 1979, Surface forces in adsorbed multilayers of water on quartz: Journal Colloid Interface Science, v. 71, p. 491-500.
Pharr, G.M. and Ashby, M.F., 1983, On creep enhanced by a liquid phase: Acta Metallurgica, Vol. 31, p. 129-138.
Porter, E.W. and James, W.C., 1986, Influence of pressure, salinity, temperature and grain size on silica diagenesis in quartzose sandstones: Chemical Geology, Vol. 57, p. 359-369.
Posey-Dowty, J., Crerar, D., Hellmann, R., and Chang, C.D., 1986, Kinetics of mineral-water reactions: theory, design and application of circulating hydrothermal equipment: American Mineralogist, v. 71, p. 85-94.
Raj, R., 1982, Creep in polycrystalline aggregates by matter transport through a liquid phase: Journal of Geophysical Research, Vol. 87, p. 4731-4739
Reuschle, T., Trotignon, L., and Guegen, Y., 1988, Pore shape evolution by solution transfer: thermodynamics and mechanics, Geophysics Journal, Vol. 95, p. 535-547
Rimstidt, J. D. and Barnes, H. L., 1980, The kinetics of silica-water reactions: Geochimica et Cosmochimica Acta, v. 44, p. 1683-1699.
Rutter, E. H., 1983, Pressure solution in nature, theory and experiment: Journal of the Geological Society of London, v. 140, p. 725-740.
Rutter, E.H., and Mainprice, D.H., 1978, The effect of water on the stress relaxation of faulted and unfaulted sandstone: Pure and Applied Geophysics, v. 116, p. 634-654
Schutjens, P.M.T.M. (1991) Experimental compaction of quartz sand at low effective stress and temperature conditions: Journal of the Geological Society of London, Vol. 148, pp. 527-539.
Spiers, C.J. and Schutjens, P.M.T.M., 1990, Densification of crystalline aggregates by fluid-phase diffusional creep, in *Deformation Processes in Minerals, Ceramics and Rocks*, edited by D.J. Barber and P.G. Meredith, p. 334-353, Unwin-Hyman, Boston.
Tada, R., Maliva, R., and Siever, R., 1987, A new mechanism for pressure solution in porous quartzose sandstone: Geochemica et Cosmochemica Acta., v. 51, p. 2295-2301.
Tada, R. and Siever, R., 1986, Experimental knife-edge pressure solution of halite, Geochimica et Cosmochimica Acta, Vol. 50, p. 29-36.
Tada, R. and Siever, R., 1989, Pressure solution during diagenesis: Annual Review Earth Planet Science, v. 17, p. 89-118.
Watson, E.B., and Brenan, J.M., 1987, Fluids in the lithosphere, part I: experimentally-determined wetting characteristics of $CO_2$-$H_2O$ fluids and their implications for fluid transport, host-rock physical properties, and fluid inclusion formation: Earth and Planetary Science Letters, v. 85, p. 497-515.
Weyl, P. K., 1959, Pressure solution and force of crystallization - a phenomenological theory: Journal of Geophysical Research, v. 64, p. 2001-2025.
Wintsch, R.P. and Dunning, J. (1985) The effect of dislocation density on the aqueous solubility of quartz and geologic implications: A theoretical approach: Journal of Geophysical Research, Vol. 90, pp. 3649-3657.
Zhang, J. Wong, T.-F., and Davis, D.M. (1990) Micromechanics of pressure-induced grain crushing in porous rocks: Journal of Geophysical Research, Vol. 95, pp. 344-352.
Zimmerman, R.W., Somerton, W.H., and King, M.S., 1986, Compressibility of porous rocks: Journal Geophysical Research, v. 91, p. 12,765-12,777.

## Symbols

| | |
|---|---|
| $A$ | pre-exponential factor in empirical creep law |
| $A_i$ | surface area of site i ($cm^2$) (i = f denotes free faces; i = c denotes contacts) |
| $a_{SiO2}$ | activity of silica in quartz at system conditions; defined to be unity at conditions of temperature and fluid pressure |
| $c$ | concentration of aqueous silica ($mol\text{-}cm^{-3}$) |
| $c_{ss}$ | steady-state silica concentration ($mol\text{-}cm^{-3}$) |
| $c^{eq}$ | equilibrium concentration of aqueous silica at conditions of temperature and fluid pressure ($mol\text{-}cm^{-3}$) |
| $E_a$ | apparent activation energy ($kj\text{-}mol^{-1}$) |
| $f$ | stress-magnification factor; translates macroscopically applied effective pressure into locally averaged normal stress at grain contacts (dimensionless) |
| $G_i$ | dissolution velocity for site i in $cm\text{-}s^{-1}$; (i = f denotes free faces; i = c denotes contacts) |
| $\Delta G$ | molar free energy for quartz-water reaction ($J\text{-}mol^{-1}$) |
| $h$ | strain-hardening coefficient |
| $J$ | flux of aqueous silica ($mol\text{-}cm^{-2}\text{-}s^{-1}$) along a reaction column |
| $k_c$ | first-order reaction rate constant for grain contacts ($cm\text{-}s^{-1}$) |
| $k_f$ | first-order reaction rate constant for grain free faces ($cm\text{-}s^{-1}$) |
| $k^+$ | dissolution rate constant for quartz in water at conditions of fluid pressure and temperature ($cm\text{-}s^{-1}$) |
| $k^-$ | precipitation rate constant for quartz in water at conditions of fluid pressure and temperature ($cm\text{-}s^{-1}$) |
| $L$ | grain size; a subscripted 'o' refers to initial or unstrained value |
| $L$ | length of reaction column (cm) |
| $n$ | number of quartz grains per unit volume ($cm^{-3}$) |
| $P_c$ | applied confining pressure (MPa) |
| $P_p$ | fluid pressure (MPa) |
| $P_e$ | effective pressure (equals the difference between $P_c$ and $P_p$) |
| $Q$ | volumetric flow rate of fluid through reaction column ($cm^3\text{-}s^{-1}$) |
| $R$ | gas constant (equals 8.314241 $MPa\text{-}cm^3\text{-}mol^{-1}\text{-}°K^{-1}$) |
| $R^+$ | rate of forward (dissolution) reaction in $mol\text{-}cm^{2}\text{-}s^{-1}$; subscript denotes effective pressure value |
| $R^-$ | rate of backward (precipitation) reaction in $mol\text{-}cm^{2}\text{-}s^{-1}$ |
| $T$ | temperature (°C or °K) |
| $t$ | time (s) |
| $V_b$ | bulk sample volume |
| $V_g$ | average grain volume of quartz sand grain ($cm^3$) |
| $V_p$ | pore volume of reaction column ($cm^3$) |
| $v$ | fluid flow velocity ($cm\text{-}s^{-1}$) |
| $u$ | fluid flow velocity during high flow rate event ($cm\text{-}s^{-1}$) |
| $\alpha,\beta$ | parameters used in empirical curve fit; see text for theoretical interpretation |
| $\phi$ | porosity |
| $\varepsilon$ | volume strain |
| $\dot{\varepsilon}$ | volume strain rate ($s^{-1}$) |
| $\nu$ | stoichiometric coefficient for quartz-water reaction (number of moles of silica released per mol of quartz dissolved) |
| $\rho$ | molar density of quartz ($mol\text{-}cm^{-3}$) |
| $\tau$ | residence time of fluid packet within reaction column (s); equal to $V_p\text{-}Q^{-1}$ or $L\text{-}v^{-1}$ for ideal plug flow |

Table 1. Thermodynamic driving forces for solution-transfer*

| Mechanism | %Increase in $SiO_2$ activity |
|---|---|
| Elastic strain energy | .001 |
| Dislocation strain energy | .050 |
| Surface Curvature (microgranulation) | 10.0 |
| Loaded Faces (normal stress) | 10.0 |

*adapted from Tada et al., 1987; T = 300°K, $P_e$ = 30 MPa

Table 2. Solid starting materials and BET surface areas

| | Size (mm) | Surface Area ($cm^2/g$) |
|---|---|---|
| Very-fine grained | .09-.12 | 878 |
| Fine grained (batch 1) | .18-.25 | 669 |
| Fine grained (batch 2) | .18-.25 | 531 |
| Medium grained | .25-.35 | 468 |

Table 3: Comparison of strain rates calculated from dissolution rates to those obtained from measured pore volume loss

| Experiment | $R^{+}_{34.5} - R^{+}_{0}$ | $\dot{\varepsilon}_v$ from dissolution kinetics ($s^{-1}$)¤ | $\dot{\varepsilon}_v$ from pore-volume loss ($s^{-1}$) |
|---|---|---|---|
| Very-fine grained sand, 150°C and 34.5 $P_e$ | 1.9E-07 | 1.3E-7 | 4.0E-8 |
| Fine grained sand, 150°C and 34.5 $P_e$ | 1.2E-07 | 4.3E-8 | 1.0E-8 |
| Fine grained sand, 200°C and 34.5 $P_e$ | 9.4E-07 | 2.8E-7 | 8.3E-8 |

¤calculated using equation (25) in text
(total grain contact area estimated from SEM micrographs is 79 $\mu m^2$ for very-fine sand and 620 $\mu m^2$ for fine-grained sand; grain size for v-fine and fine grained sand is taken to be .011 and .022 cm respectively)

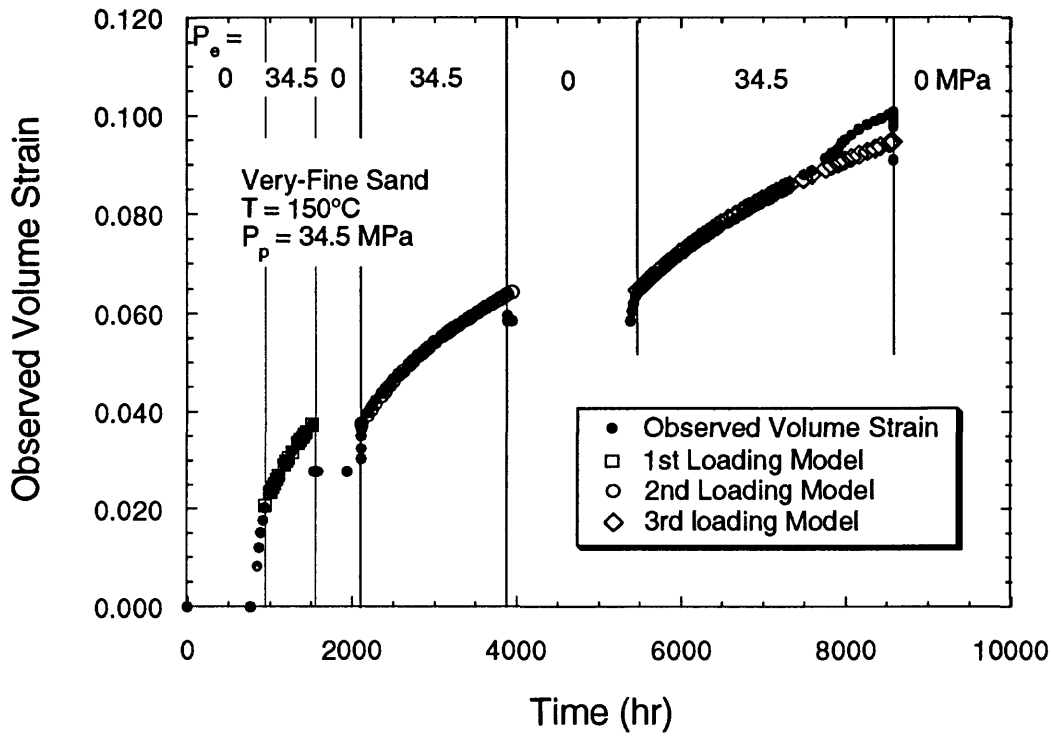



Figure 1. Changes in volume strain with time for very-fine grained St. Peter sand subject to three episodes of loading to 34.5 MPa effective pressure. Prior to the introduction of an effective pressure, the sand was exposed to hydrostatic pressure of 34.5 MPa and 150°C for ~800 hours, whereupon confining pressure was increased at a rate of 6.9 MPa per day for five days, with pore pressure kept constant. Volume strain is determined from the volume of pore fluid extracted to keep a constant effective pressure on the sample. During each loading episode (with effective pressure of 34.5 MPa), the rate of volume strain accumulation slows with time (or with increasing volume strain). Also shown are least-squares regression fits of a model equation, discussed in the text, for the strain-time behavior. The large deviation of the model from the observed strain during the third and final loading phase is due to a malfunction in the temperature controlling system, which led to a 2-3° temperature increase for a few days. We neglected these data in our modeling considerations.

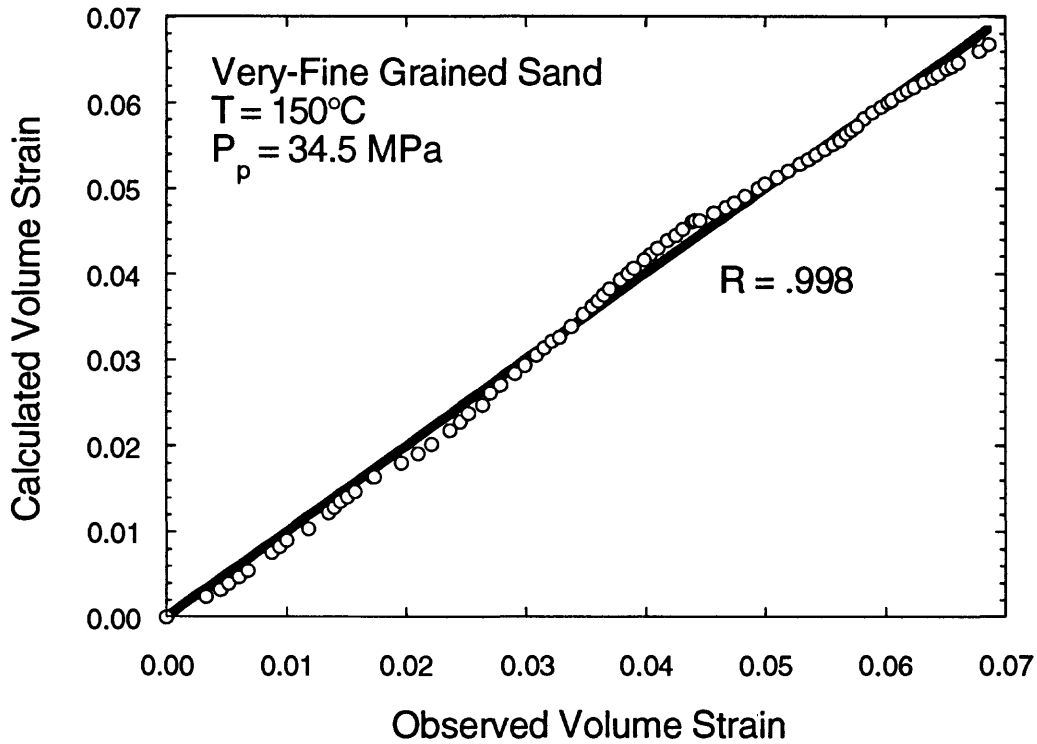



Figure 2. Comparison of observed volume strain in the experiment described in Figure 1 to that predicted from the proposed strain-time model. The agreement between observed and predicted strains is excellent.

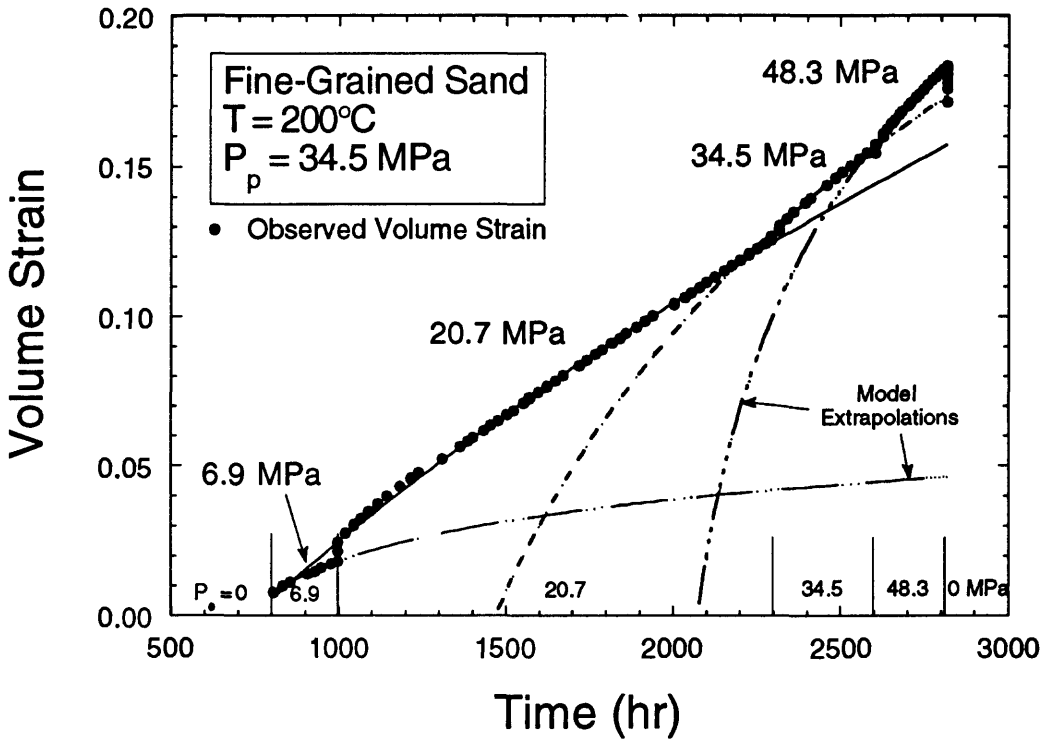



Figure 3. Volume strain plotted versus time for an experiment using fine-grained St. Peter sand subject to 200°C, 34.5 MPa fluid pressure, and effective pressure ranging from 34.5 to 82.8 MPa. Here the effective pressure was increased in a step-wise fashion. Also depicted by curves are model extrapolations back to zero strain using our strain-time model. The "limiting strain rate" discussed in the text is the slope of these lines as the volume strain approaches zero.

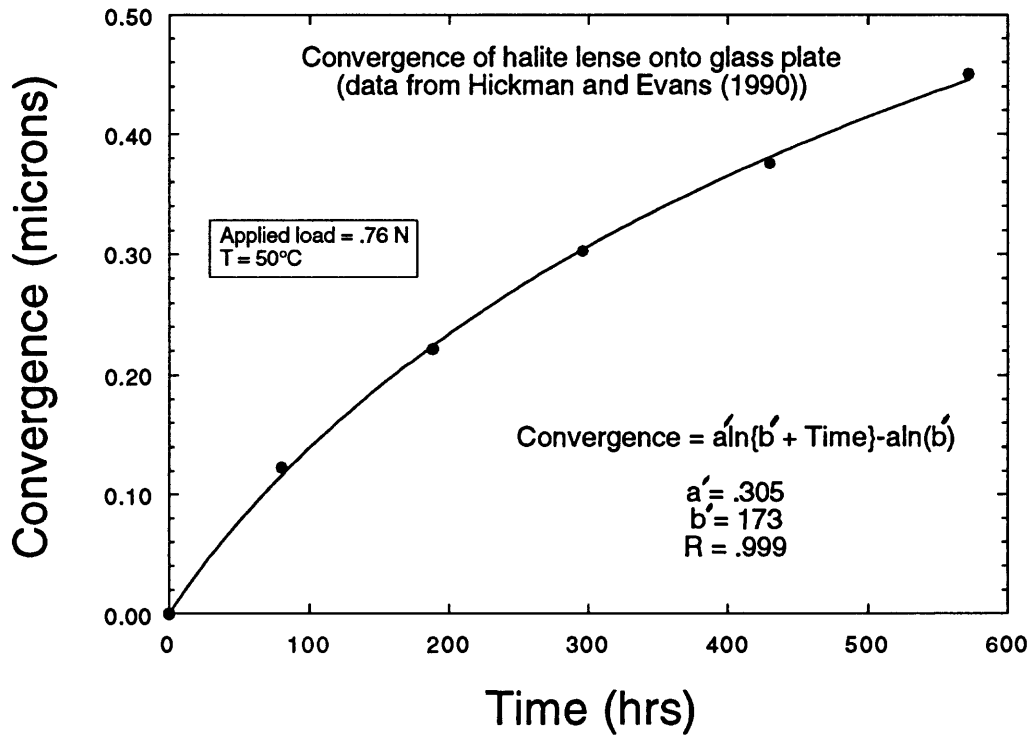



Figure 4. Increasing convergence with time of a polished halite "lens" against a flat glass plate (from Hickman and Evans, 1990), due presumably to stress-induced solution-transfer from the portions of the halite lens in contact with glass. The rate of convergence slows with time (or with increasing convergence); we have used an equation similar to that used in Figures 1 through 3 to fit the observed convergence-time data. Because convergence of grains at a single contact may be related to aggregate volume strain (see text for details) involving many grain contacts, the similarity of the "strain-hardening" relationship seen in both our experiments and those of Hickman and Evans (1990) suggest that the mechanism for how strain slows with increasing strain is the same in both cases, namely the broadening of grain contact areas. This increase in contact area, in the quartz sand experiments, may be also due to a stress-induced solution transfer.

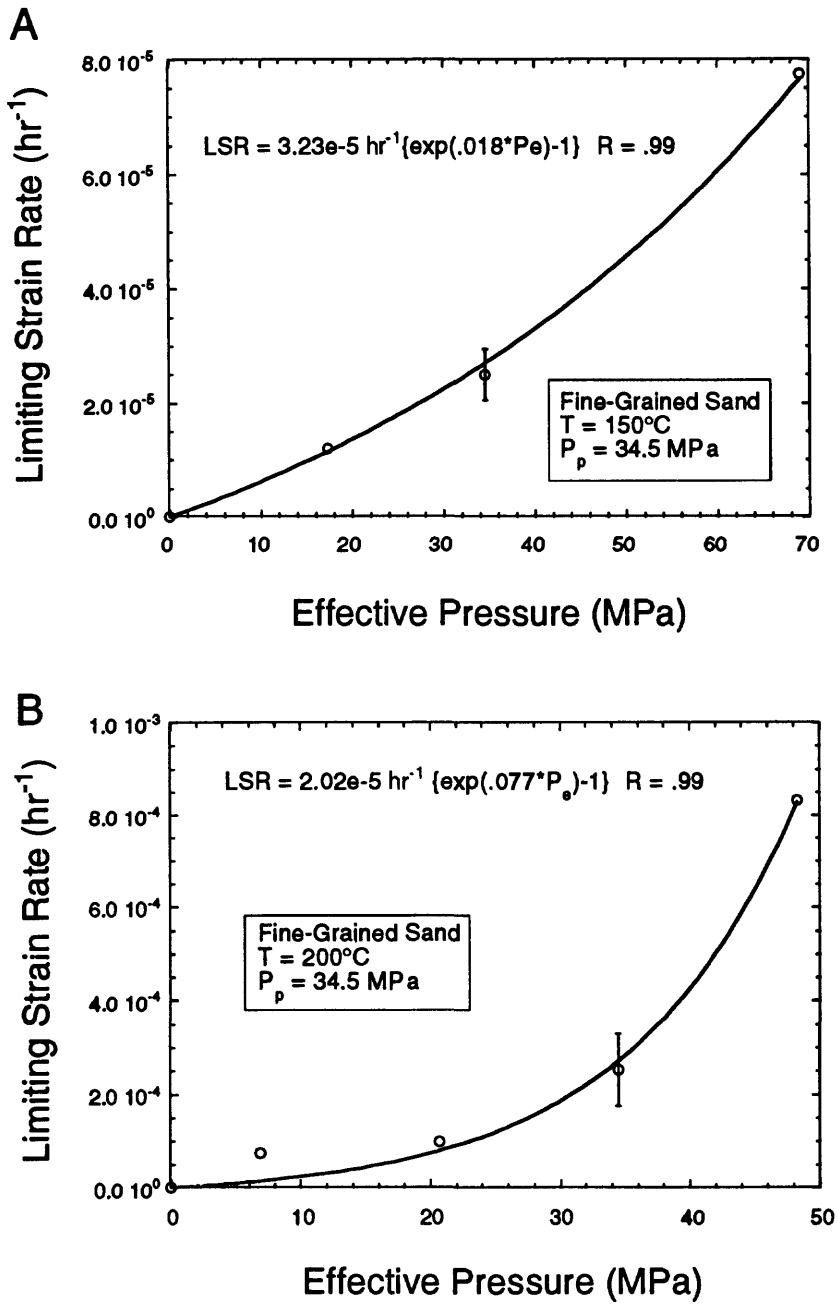



Figure 5. A compilation of several experiments using fine-grained St. Peter sand at (a) 150°C and (b) 200°C and 34.5 MPa pore pressure, plotting the limiting strain rate as a function of effective pressure. The limiting strain rate was determined in each case from a least-squares regression analysis of the raw strain-time data using a procedure outlined in the text. A simple mathematical expression involving the exponential of effective pressure fits the data sets at both temperatures to an excellent degree. The error bars shown are twice the standard deviation calculated from results of three separate experiments run at the same conditions, and so are a measure of reproducibility.

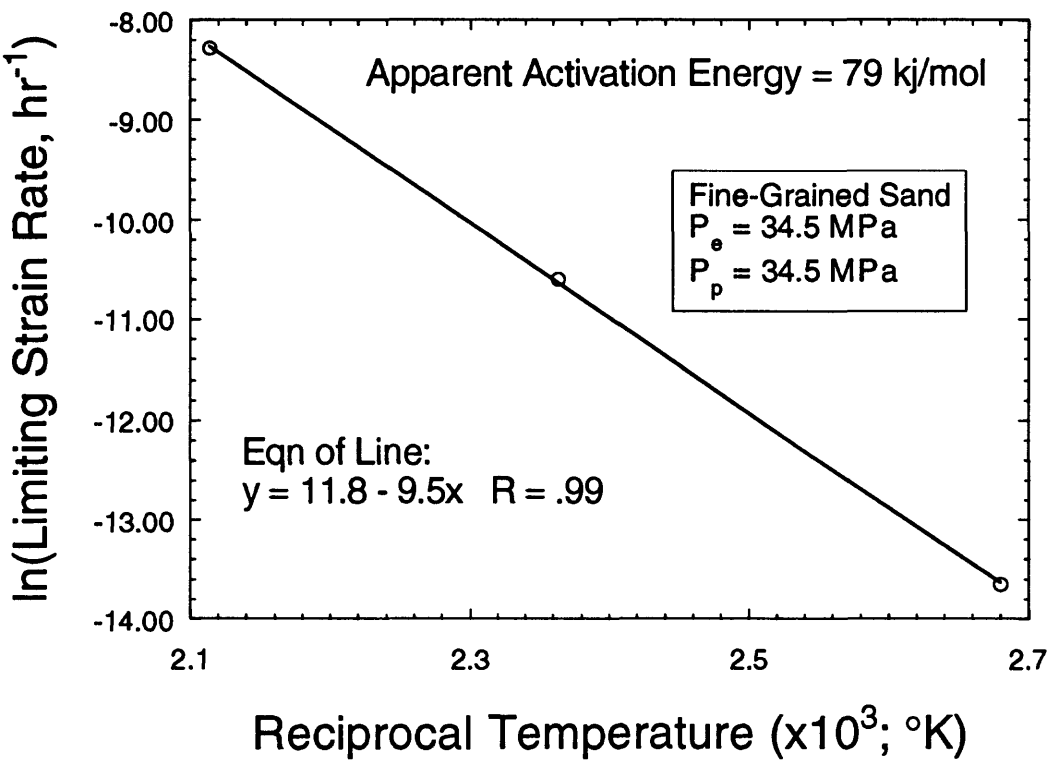



Figure 6. Relationship between limiting strain rate and temperature, determined from data sets at 100°, 150°, and 200°C. The linear fit of the data in the space of the natural logarithm of the limiting strain rate and reciprocal temperature defines a typical Arrhenius relationship, with an effective activation energy of 79 kj/mol determined from the slope of the curve fit.

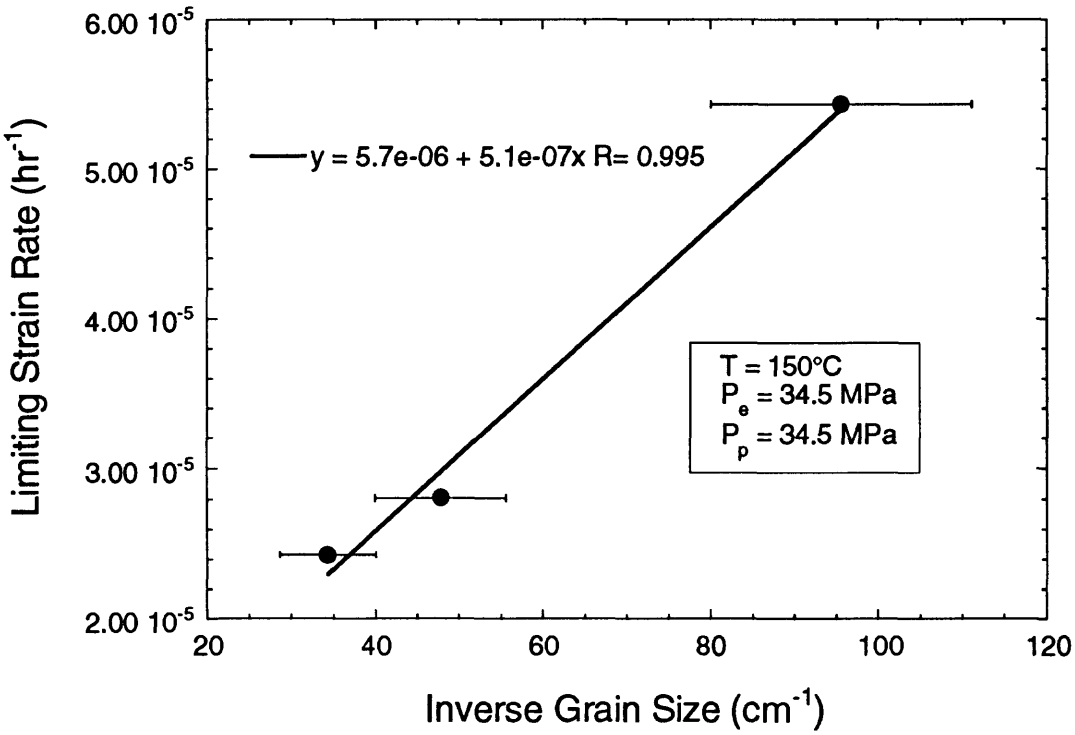



Figure 7. This plot shows the grain size effect on the limiting strain rate for three sets of experiments using very-fine, fine, and medium grained St. Peter sand run at 34.5 MPa effective and pore pressures and 150°C. The range in grain size (horizontal error bars) defines the range in mesh sizes used to originally size the sand (.09-.12, .18-.25, and .25 to .35 mm). A linear fit of the limiting strain rate to the inverse of grain size fits the data very well, although other dependencies are possible, and we have investigated only three grain sizes.

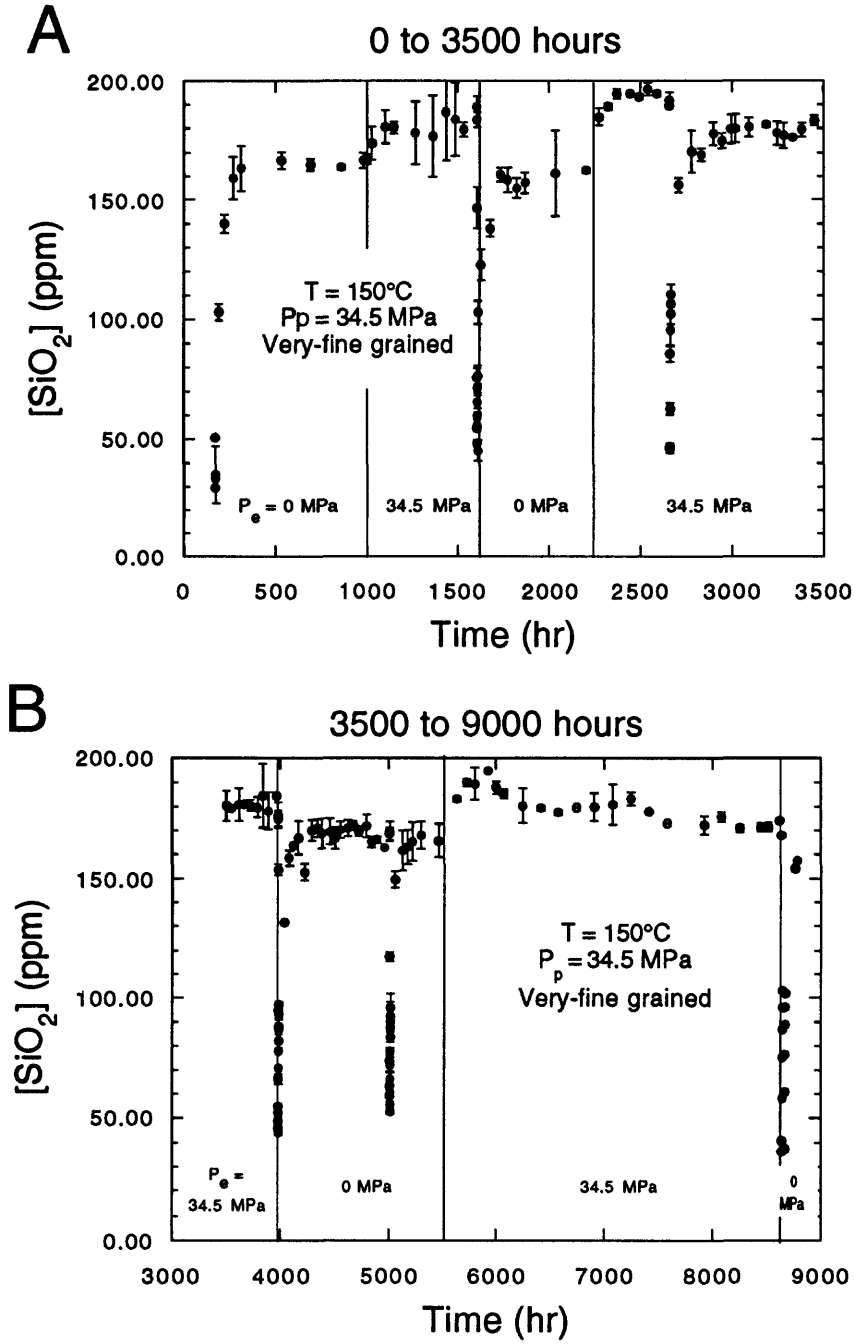



Figure 8. Behavior of dissolved aqueous silica with time for the experiment depicted in Figure 1, using very-fine grained sand in distilled water at 150°C and 34.5 MPa pore pressure. The concentration at steady-state (and independent of flow rate) increases during the loading episodes of effective pressure equal to 34.5 MPa but returns to the background level (quartz saturation in distilled water at these temperature and fluid pressure conditions) upon removal of the effective pressure. It is also evident, especially in the third loading episode, that the silica concentration decreases with time, or alternatively with increasing volume strain. The abrupt downward "spikes" of concentration that are evident periodically throughout the experiment are times when the fluid flow rate was increased in order to examine changes in dissolution kinetics during the course of the experiment.

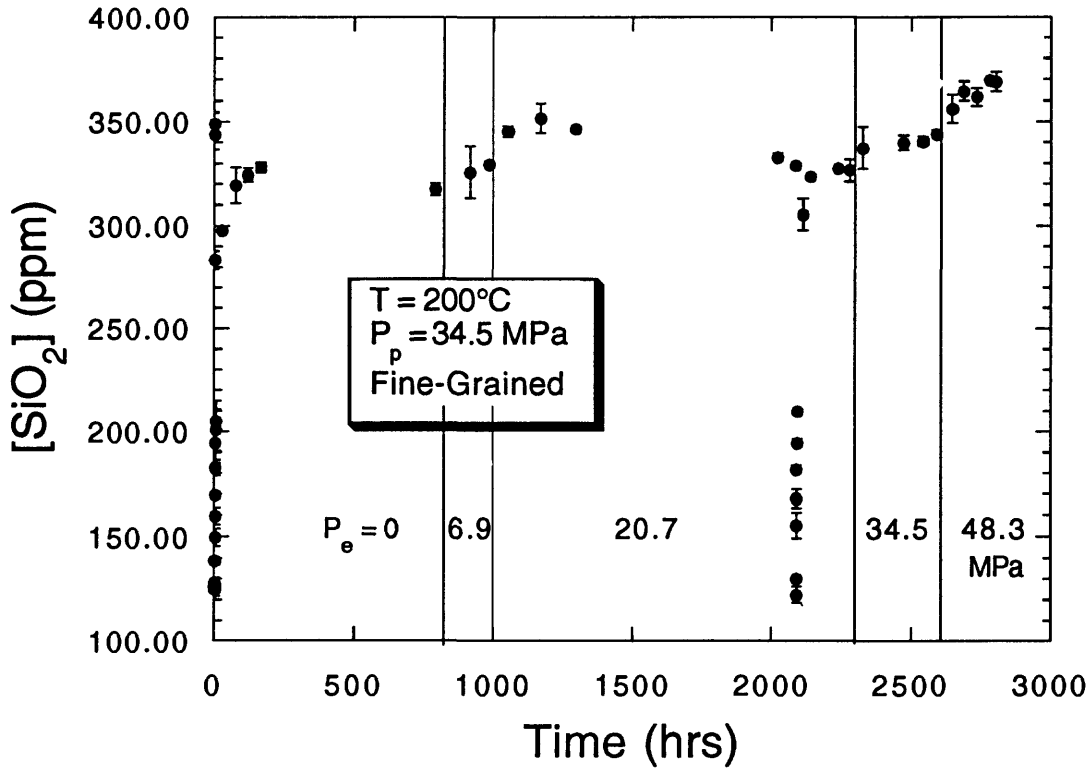
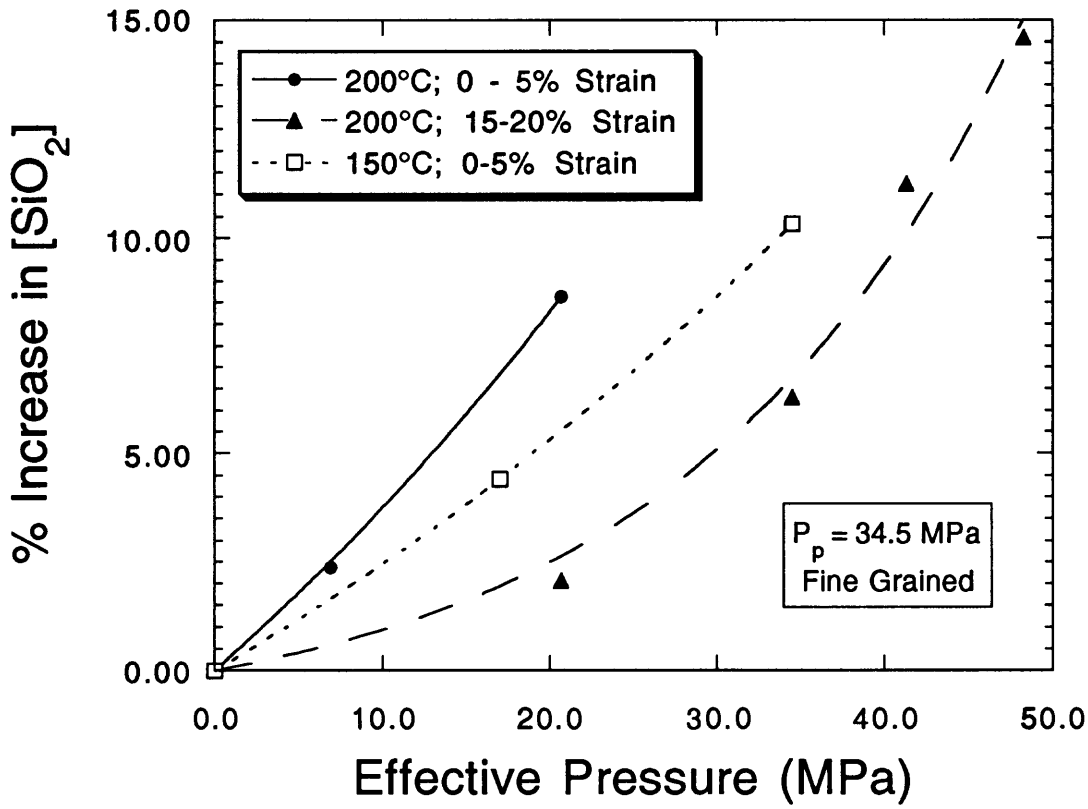



Figure 9. Silica concentration versus time for the experiment discussed in Figure 3. The concentration increases with rising effective pressure at nearly constant volume strain, but decreases with increasing volume strain at constant effective pressure. The change in concentration with varying volume strain may reflect a broadening in grain contact area, and a concomitant decrease in the normal stress across grain contacts.




| % Increase = 100x{exp(α*Pe)-1}/(β+1) | α | β |
|---|---|---|
| 200°C; 0-5% strain | .019 | 4.7 |
| 200°C; 15-20% strain | .050 | 69.9 |
| 150°C; 0-5% strain | .016 | 5.9 |

Figure 10. Percent increase in steady-state silica concentration with increasing effective pressure, for three sets of experiments using fine-grained sand at nearly constant volume strain (150°C, 0-5% strain; 200°C, 0-5% strain; and 200°C, 15-20% strain). An equation equating the % increase with the exponential of effective pressure fits all data sets to an excellent degree of correlation.

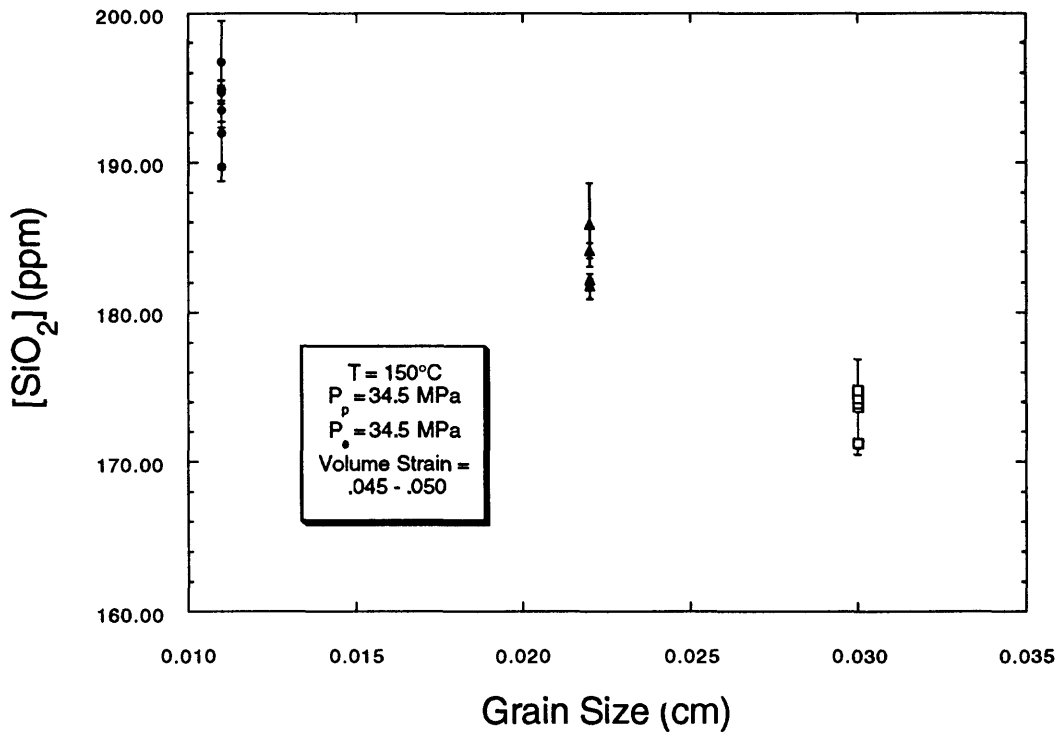



Figure 11. Effect of grain size on the steady-state silica concentration at low (0-5%) volume strains, 150°C, and 34.5 MPa effective pressure. The (flow-rate independent) concentration rises as grain size decreases; we argue that this dependence arises from the slight lowering of quartz precipitation rate per single grain as the grain size is decreased. St. Peter quartz sand saturation in distilled water at conditions of 150°C and 34.5 MPa pressure corresponds to 160 ppm, coinciding with the x-axis.

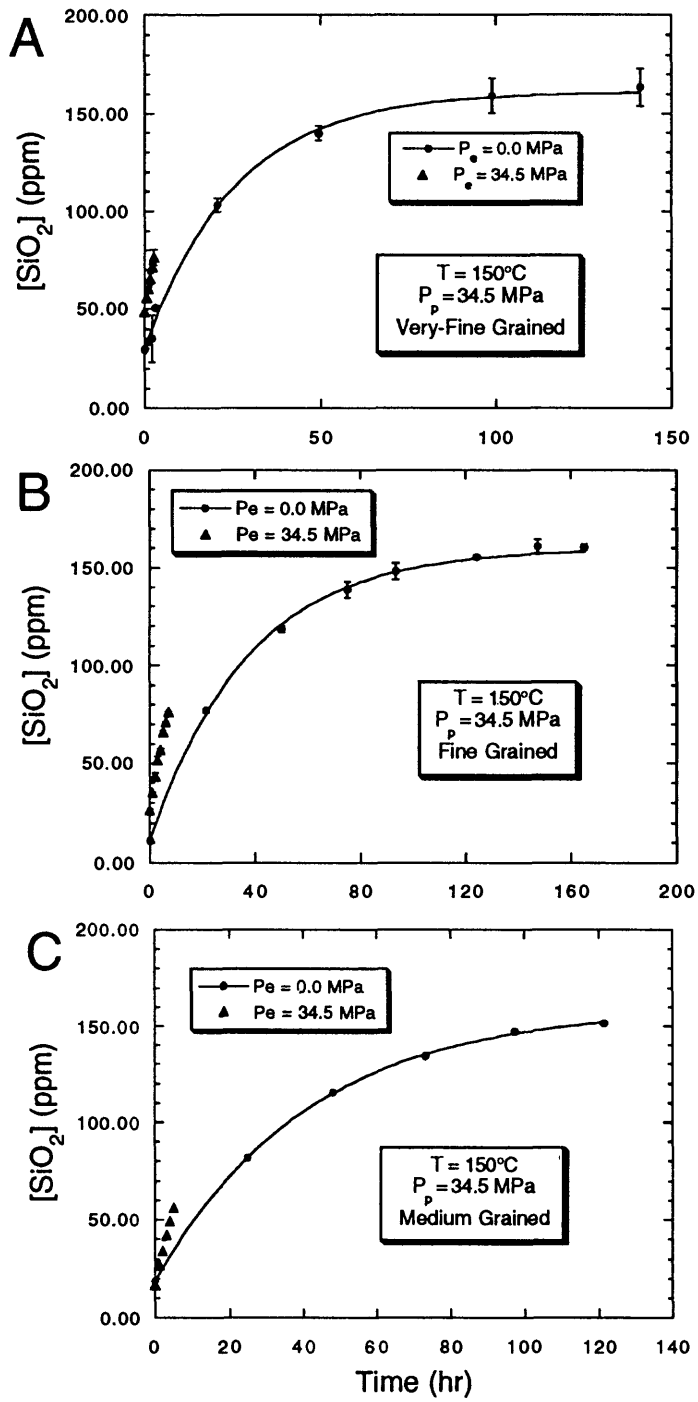



Figure 12. Three studies examining the effect of grain size and effective pressure on the sand dissolution rate in distilled water. Shown are the dissolved silica with time for (a) very-fine, (b) fine, and (c) medium grained sand at 150°C, 0.0 and 34.5 MPa effective pressure, and < 5% volume strain. The solid lines fitting the 0.0MPa data sets are from a first-order kinetics solution for quartz sand in a column reactor undergoing perfect plug-flow. As the slope of the lines fitting the concentration-time data sets are roughly proportional to the dissolution rate, it is evident that the sand dissolves faster when exposed to an effective pressure compared to the rate at zero effective pressure.

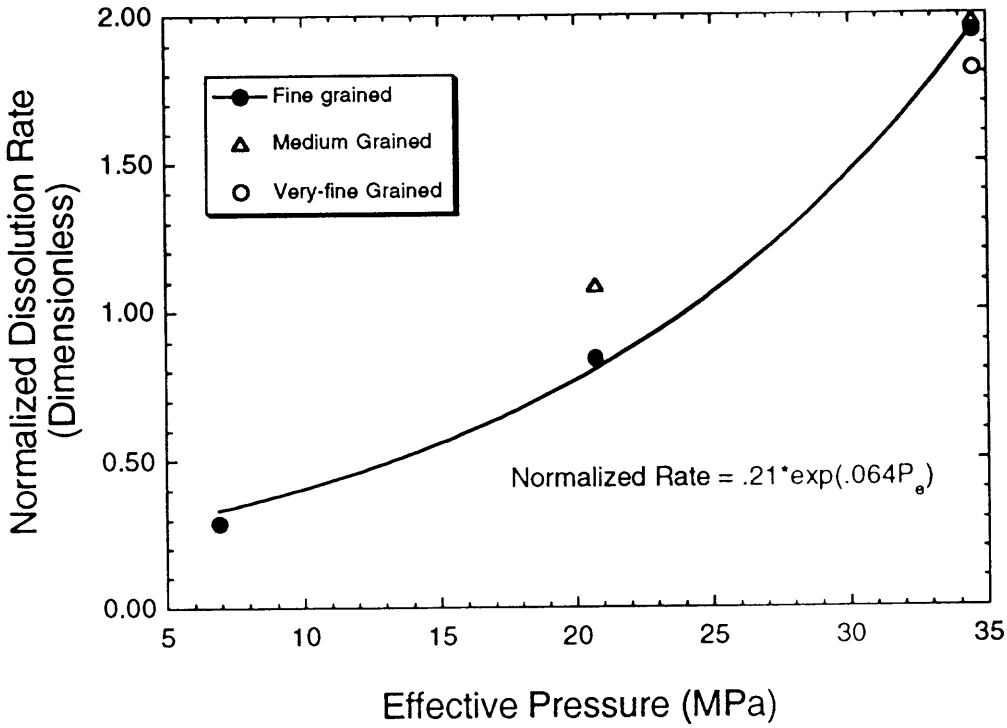
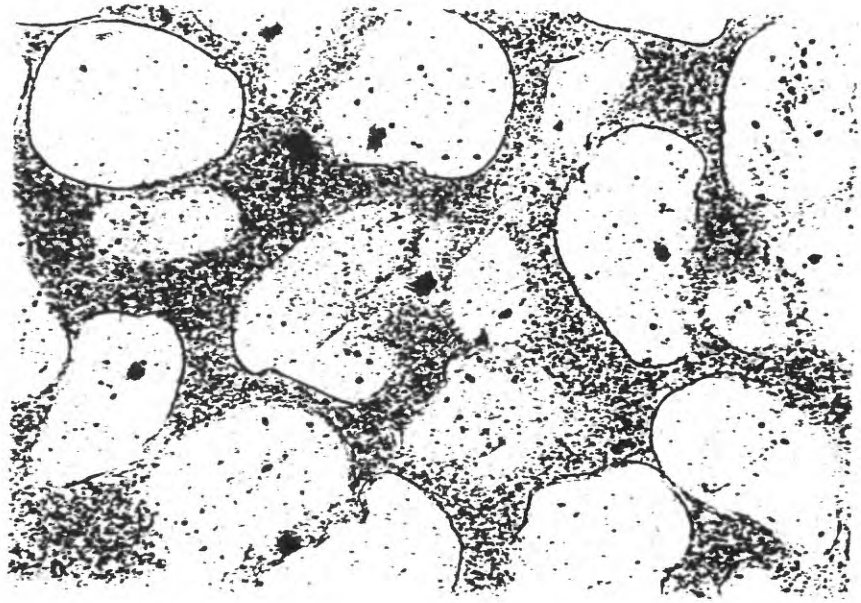
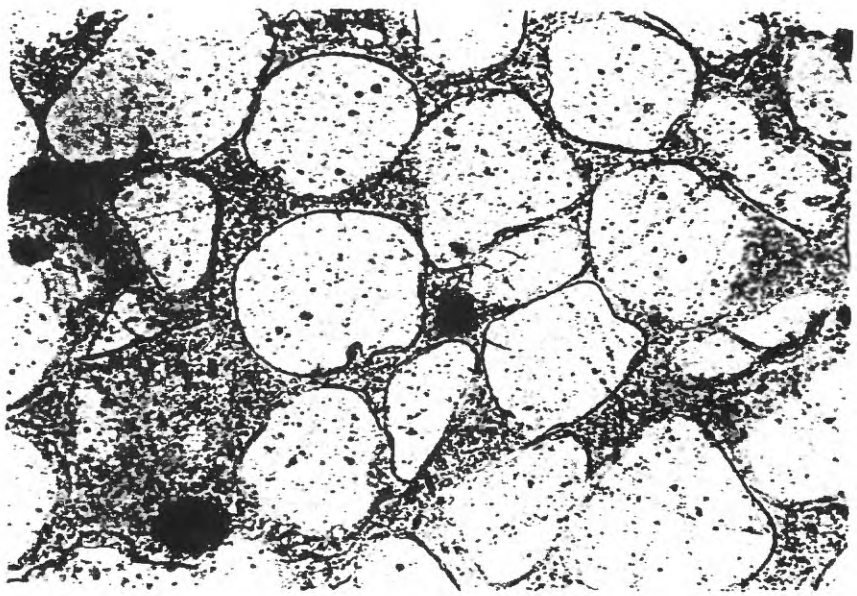



Figure 13. Normalized dissolution rate as a function of effective pressure for sands of different grain sizes at 150°C. The data fall roughly along the same trend, suggesting that when background effects (i.e. dissolution at zero effective pressure) are subtracted out, the effective-pressure dependent dissolution rate (that occurring, perhaps, at grain contacts) does not change with grain size.

A



B



C

Figure 14. These photomicrographs depict fine-grained sand with initial porosity of ~35% from our experimental run products that have been subject to three different conditions. (a) Sand that has undergone loading to 34.5 MPa effective pressure and 150°C, whereupon the experiment was terminated. Grains appear to "float" in epoxy. (b) Sand exposed to 34.5 MPa effective pressure and 200°C for nearly three thousand hours, which was sufficient to reduce porosity to below 20%. Note the abundant fluid inclusion trails, which appear to be more numerous than in (a), and many grain-grain contacts. (c) Sand subject to effective pressures above 55.2 MPa, which at 200°C was sufficient to induce grain crushing. The epoxy was injected while the samples were still jacketed and water-wet; this accounts for the less-than perfect appearance of the epoxy matrix and likely forced many grain contacts apart.

## CONFERENCES TO DATE

| | |
|---|---|
| Conference I | Abnormal Animal Behavior Prior to Earthquakes, 1<br>Not Open-Filed |
| Conference II | Experimental Studies of Rock Friction with Application to Earthquake Prediction<br>Not Open-Filed |
| Conference III | Fault Mechanics and Its Relation to Earthquake Predition<br>Open-File No. 78-380 |
| Conference IV | Use of Volunteers in the Earthquake Hazards Reduction Program<br>Open-File No. 78-336 |
| Conference V | Communicating Earthquake Hazard Reduction Information<br>Open-File No. 78-933 |
| Conference VI | Methodology for Identifying Seismic Gaps and Soon-to-Break Gaps<br>Open-File No. 78-943 |
| Conference VII | Stress amd Strain Measurements Related to Earthquake Prediction<br>Open-File No. 79-370 |
| Conference VIII | Analysis of Actual Fault Zones in Bedrock<br>Open-File No. 79-1239 |
| Conference IX | Magnitude of Deviatoric Stresses in the Earth's Crust and Upper Mantle<br>Open-File No. 80-625 |
| Conference X | Earthquake Hazards Along the Wasatch and Sierra-Nevada Frontal Fault Zones<br>Open-File No. 80-801 |
| Conference XI | Abnormal Animal Behavior Prior to Earthquakes, II<br>Open-File No. 80-453 |
| Conference XII | Earthquake Prediction Information<br>Open-File No. 80-843 |
| Conference XIII | Evaluation of Regional Seismic Hazards and Risk<br>Open-File No. 81-437 |
| Conference XIV | Earthquake Hazards of the Puget Sound Region, Washington<br>Open-File No. 82-19 |
| Conference XV | A Workshop on "Preparing for and Responding to a Damaging Earthquake in the Eastern United States"<br>Open-File No. 82-220 |
| Conference XVI | The Dynamic Characteristics of Faulting Inferred from Recording of Strong Ground Motion<br>Open-File No. 82-591 |
| Conference XVII | Hydraulic Fracturing Stress Measurements<br>Open-File No. 82-1075 |
| Conference XVIII | A Workshop on " Continuing Actions to Reduce Losses from Earthquakes in the Mississippi Valley Area"<br>Open-File No. 83-157 |

| | |
|---|---|
| Conference XIX | Active Tectonic and Magmatic Processes Beneath Long Valley<br>Open-File No. 84-939 |
| Conference XX | A Workshop on "The 1886 Charleston, South Carolina, Earthquake and its Implications for Today"<br>Open-File No. 83-843 |
| Conference XXI | A Workshop on "Continuing Actions to Reduce Potential Losses from Future Earthquakes in the Northeastern United States"<br>Open-File No. 83-844 |
| Conference XXII | A Workshop on "Site-Specific Effects of Soil and Rock on Ground Motion and the Implications for Earthquake-Resistant Design"<br>Open-File No. 83-845 |
| Conference XXIII | A Workshop on "Continuing Actions to Reduce Potential Losses from Future Earthquakes in Arkansas and Nearby States"<br>Open-File No. 83-846 |
| Conference XXIV | A Workshop on "Geologic Hazards in Puerto Rico"<br>Open-File No. 84-761 |
| Conference XXV | A Workshop on "Earthquake Hazards in the Virgin Islands Region"<br>Open-File No. 84-762 |
| Conference XXVI | A Workshop on "Evaluation of the Regional and Urban Earthquake Hazards in Utah"<br>Open-File No. 84-763 |
| Conference XXVII | Mechanics of the May 2, 1983 Coalinga Earthquake<br>Open-File No. 85-44 |
| Conference XXVIII | A Workshop on "The Borah Peak, Idaho, Earthquake"<br>Open-File No. 85-290 |
| Conference XXIX | A Workshop on "Continuing Actions to Reduce Potential Losses from Future Earthquakes in New York and Nearby States"<br>Open-File No. 85-386 |
| Conference XXX | A Workshop on "Reducing Potential Losses from Earthquake Hazards in Puerto Rico"<br>Open-File No. 85-731 |
| Conference XXXI | A Workshop on "Evaluation of Regional and Urban Earthquake Hazards and Risk in Alaska"<br>Open-File No. 86-79 |
| Conference XXXII | A Workshop on "Future Directions in Evaluating Earthquake Hazards of Southern California"<br>Open-File No. 86-401 |
| Conference XXXIII | A Workshop on "Earthquake Hazards in the Puget Sound, Washington Area"<br>Open-File No. 86-253 |
| Conference XXXIV | A Workshop on "Probabilistic Earthquake-Hazards Assessments"<br>Open-File No. 86-185 |

| | |
|---|---|
| Conference XXXV | A Workshop on "Earth Science Considerations for Earthquake Hazards Reduction in the Central United States"<br>Open-File No. 86-425 |
| Conference XXXVI | A Workshop on "Assessment of Geologic Hazards and Risk in Puerto Rico"<br>Open-File No. 87-007 |
| Conference XXXVII | A Workshop on "Earthquake Hazards Along the Wasatch Fault, Utah"<br>Open-File No. 87-154 |
| Conference XXXVIII | A Workshop on "Physical & Observational Basis for Intermediate Term Earthquake Prediction"<br>Open-File 87-154 |
| Conference XXXIX | Directions in Paeloseismology<br>Open-File No. 87-673 |
| Conference XL | A Workshop on "The U.S. Geological Survey's Role in Hazards Warnings"<br>Open File No. 87-269 |
| Conference XLI | A Review of the Earthquake Research Applications in the National Earthquake Hazard Reduction Program: 1977-1987<br>Open-File No. 88-13-A |
| Conference XLII | A Workshop on "Evaluation of Earthquake Hazards and Risk in the Puget Sound and Portland Areas<br>Open File No. 88-541 |
| Conference XLIII | A Workshop on "Earthquake Risk: Information Needs of the Insurance Industry"<br>Open-File No. 88-669 |
| Conference XLIV | A Workshop on "Geological, Geoophysical, and Tectonic Settings of the Cascade Range"<br>Open-File No. 89-178 |
| Conference XLV | A Workshop on "Fault Segmentation and Controls of Rupture Initiation and Termination"<br>Open-File No. 89-315 |
| Conference XLVI | The 7th U.S.- Japan Seminar on Earthquake Prediction<br>Open-File No. 90-98 |
| Conference XLVII | A Workshop on "USGS'S New Generation of Probabilistic Ground Motion Maps and their Applications to Building Codes"<br>Open-File No. 89-364 |
| Conference XLVIII | A Workshop on "Earthquake Hazards in the Puget Sound, Portland Area"<br>Open-File No. 89-465 |
| Conference IL | A Meeting of the U.S. Ad HOC Working Group on: Earthquake Related Casualties"<br>Open-File No. 90-244 |
| Conference LX | 4th Annual Workshop on "Earthquake Hazards in the Puget Sound, Portland Area"<br>Open-File No. 90-703 |
| Conference LXI | A Workshop on "Role of State Geological Surveys in Postearthquake Investigations"<br>Open-File 93-346 |

| | |
|---|---|
| Conference LXII | Proceeding of the "Eighth Joint Meeting of the U.S.-Japan Conference on Natural Resources (UJNP) Open-File No. 93-542 |
| Conference LXIII | USGS Red-Book Conference on "The Mechanical Involement of Fluids in Faulting" Open-File No. 94-228 |

For information on ordering the above publications, please contact:
U.S. Geological Survey
Books and Open-File Reports Service Section
Building 41, Box 25425
Federal Center
Denver, Colorado 80225



* U.S. GOVERNMENT PRINTING OFFICE: 1994-585-853/19200