THE MECHANICS OF EARTHQUAKES AND FAULTING

Third Edition

Christopher H. Scholz

The Mechanics of Earthquakes and Faulting

3rd Edition

This essential reference for graduate students and researchers provides a unified treatment of earthquakes and faulting as two aspects of brittle tectonics at different timescales. The intimate connection between the two is manifested in their scaling laws and populations, which evolve from fracture growth and interactions between fractures. The connection between faults and the seismicity generated is governed by the rate- and state-dependent friction laws – producing distinctive seismic styles of faulting and a gamut of earthquake phenomena, including aftershocks, afterslip, earthquake triggering, and slow slip events. The third edition of this classic treatise presents a wealth of new topics and new observations. These include slow earthquake phenomena; friction of phyllosilicates and at high sliding velocities; fault structures; relative roles of strong and seismogenic versus weak and creeping faults; dynamic triggering of earthquakes; oceanic earthquakes; megathrust earthquakes in subduction zones; deep earthquakes; and new observations of earthquake precursory phenomena.

Christopher H. Scholz is an emeritus professor at Lamont Doherty Earth Observatory, Columbia University, where, over his 50-year career, he has published more than 300 papers on rock mechanics, fault mechanics, and the physics of earthquakes. He is a Fellow of the American Geophysical Union, and has been awarded the Murchison Medal by The Geological Society of London, and the Harry Fielding Reid Medal by the Seismological Society of America.

A photograph by G. K. Gilbert of the surface rupture produced by the 1906 San Francisco earthquake. (Photo courtesy of the US Geological Survey.)

THE MECHANICS OF EARTHQUAKES AND FAULTING

3rd Edition

CHRISTOPHER H. SCHOLZ
Lamont Doherty Earth Observatory, Columbia University, New York

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107163485
DOI: 10.1017/9781316681473

First published 2019
First paperback edition 1990
Reprinted 1992, 1994, 1997, 2000
Second edition 2002
Seventh printing 2012
Third edition 2019

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Scholz, C. H. (Christopher H.), author.
Title: The mechanics of earthquakes and faulting / Christopher H. Scholz (Columbia University, New York).
Description: 3rd edition. | Cambridge ; New York, NY : Cambridge University Press, 2018. | Includes bibliographical references and index.
Identifiers: LCCN 2018021900 | ISBN 9781107163485
Subjects: LCSH: Seismology. | Earthquakes. | Faults (Geology)
Classification: LCC QE534.3 .S37 2002 | DDC 551.22–dc23
LC record available at https://lccn.loc.gov/2018021900

ISBN 978-1-107-16348-5 Hardback
ISBN 978-1-316-61523-2 Paperback

. . . and then there would not be friction any more, and the sound would cease, and the dancers would stop. . .

Leonardo da Vinci

From a notebook dated September, 1508
MacCurdy, E. 1958. *The Notebooks of Leonardo da Vinci*, p. 282, New York: George Braziller

Contents

Color plate section to be found between pp. 266 and 267

Preface to the first edition

It has now been more than thirty years since the publication of E. M. Anderson's *The Dynamics of Faulting* and C. F. Richter's *Elementary Seismology*. Several generations of earth scientists were raised on these texts. Although these books are still well worth reading today for their excellent descriptions of faults and earthquakes, the mechanical principles they espoused are now well understood by the undergraduate student at the second or third year. In the meantime a great deal has been learned about these subjects, and the two topics, faulting and earthquakes, described in those books have merged into one broader field, as earthquakes have been more clearly understood to be one manifestation of faulting. During this period of rapid progress there has not been a single book written that adequately fills the gap left by these two classics. As a result it has become increasingly difficult for the student or active researcher in this area to obtain an overall grasp of the subject that is both up-to-date and comprehensive and that is based firmly on fundamental mechanical principles. This book has been written to fill this need.

Not least among the difficulties facing the researcher in this field is the interdisciplinary nature of the subject. For historical reasons earthquakes are considered to be the province of the seismologist and the study of faults is that of the geologist. However, because earthquakes are a result of an instability in faulting that is so pervasive that on many faults most slip occurs during them, the interests of these two disciplines must necessarily become intertwined. Moreover, when considering the mechanics of these processes the rock mechanicist also becomes involved, because the natural phenomena are a consequence of the material properties of the rock and its surfaces.

It is a consequence of the way in which science is organized that the scientist is trained by discipline, not by topic, and so interdisciplinary subjects such as this one tend to be attacked in a piecemeal fashion from the vantage of the different specialties that find application in studying it. This is disadvantageous because progress is hindered by lack of communication between the different disciplines, misunderstandings can abound, and different, sometimes conflicting, schools of thought can flourish in the relative isolation of separate fields. Workers in one field may be ignorant of relevant facts established in another, or, more likely, be unaware of the skein of evidence that weights the convictions of workers in another field. This leads not only to a neglect of some aspects in considering a question, but also to the quoting of results attributed to another field with greater confidence than workers in that field would themselves maintain. It is not enough to be aware, secondhand, of the contributions of another field – one must know the basis, within the internal structure of the evidence and tools of that field, upon which that result is maintained. Only then is one in a position to take the results of all the disciplines and place them, with their proper weight, in the correct position of the overall jigsaw puzzle. Because the literature on this topic has become both large and diverse, a guide is useful in this process, together with some unifying mechanical principles that allow the contours of the forest to be seen from between the trees.

Although I have dabbled, to one degree or another, in the various different disciplinary approaches to this problem and therefore have a rudimentary working knowledge of them, my own specialty is rock mechanics, and so this approach is the one most emphasized in this book. Faults are treated as shear cracks, the propagation of which may be understood through the application of fracture mechanics. The stability of this fault movement, which determines whether the faulting is seismic or aseismic, is determined by the frictional constitutive law of the fault surface, and so that is the second major theme applied throughout this treatment. The application of these principles to geology is not straightforward. One cannot actually do a laboratory experiment that duplicates natural conditions. Laboratory studies can only be used to establish physical processes and validate theories. To apply the results of this work to natural phenomena requires a conceptual jump, because of problems of scale and because both the nature of the materials and the physical conditions are not well known. In order to do this one must have constant recourse to geological and geophysical observations and, working backwards, through these physical principles determine the underlying cause of the behavior of faults. For this reason, much of this book is taken up in describing observations of natural cases.

Because rock mechanics is not taught universally in earth science curricula, the first two chapters present an account of brittle fracture and friction of rock, beginning from first principles. These chapters provide the basis for the later discussion of geological phenomena. The subsequent chapters assume a beginning graduate level understanding of the earth science disciplines involved. In these chapters the results of geology, seismology, and geodesy are presented, but the techniques employed by the various specialties are not described at any length. The emphasis is on providing an overall understanding of a scientific topic rather than teaching a specific craft. A goal was to describe each topic accurately, but at such a level that it could be understood by workers in other fields.

A book may be structured in many different ways. In this case, I found it difficult to choose between organizing the book around the physical mechanisms or around the natural phenomena in which they are manifested. The latter scheme would be more familiar to the earth scientist, the former to the mechanicist. Ultimately, I adopted a system arranged around mechanics, but which still retains many of the more familiar traditional associations. Because some mechanisms are important in a number of different phenomena, which might otherwise be considered quite distant, and some earthquakes provide examples of several phenomena, there are often more than two connections to other topics. Therefore, it was not always possible to present the subject matter in a serial sequence. I consequently adopted a system of cross-referencing that allows the reader to traverse the book in alternative paths. I hope this system will be more helpful than confusing.

When I first entered graduate school twenty-five years ago, most of the material described in this book was not yet known. The first generation of understanding, outlined in Anderson's and Richter's books, has been augmented by a second generation of mechanics, much more thorough and quantitative than the preceding. This has been a most productive era, which this book celebrates. I owe my own development to associations with many people. My first mentor, W. F. Brace, set me on this path, and the way has been lit by many others since. I have also been a beneficiary of an enlightened system of scientific funding during this period, which has allowed me to pursue many interesting topics, often at no little expense. For this I particularly would like to thank the National Science Foundation, the US Geological Survey, and NASA.

Many have helped in the preparation of this book. In particular I acknowledge the assistance of my editor, Peter–John Leone; Kazuko Nagao, who produced many of the illustrations; and those who have reviewed various parts of the manuscript: T.–F. Wong, W. Means, J. Logan, S. Das, P. Molnar, J. Boatwright, L. Sykes, D. Simpson, and C. Sammis. Particular thanks are due to T. C. Hanks, who offered many helpful comments on the text, and who, over the course of a twenty–year association, has not failed to point out my foibles. I dedicate the book to my wife, Yoshiko, who provided me with the stability in my personal life necessary for carrying out this task.

Preface to the second edition

When the first edition of this book was completed in 1989 the study of earthquakes and faulting was still developing rapidly and has continued to do so in the intervening years. It thus seemed necessary, in order to keep this work useful, that an extensively revised and updated new edition be prepared.

Progress during these dozen years has not, of course, been uniform. There have been rapid developments in some areas whereas others have been relatively static. As a result, some sections and chapters have been extensively revised while others remain almost the same, undergoing only minor updating. A goal in this revision was to retain the same overall length, and this has been largely successful. This necessitated the removal of material which in hindsight no longer seemed as vital as it once did or which had been superseded by more recent results.

The two major themes of the first edition have been further developed in the interim. The first of these is the intimate connection between fault and earthquake mechanics. Fault mechanics in 1989 was still in a primitive state, but rapid progress during the 1990s has brought the discovery of the main fault scaling laws, the nature of fault populations, and how these result from the processes of fault growth and interaction. This new knowledge of fault mechanics provides a fuller appreciation of faulting and earthquakes as two aspects of the same dynamical system: the former its long-timescale and the latter its short-timescale manifestation. One major development along these lines is the realization that neither faulting nor earthquakes behave in an isolated manner but interact with other faults or earthquakes through their stress fields, sometimes stimulating the activity of neighboring faults, sometimes inhibiting it, the totality of such interactions resulting in the populations, of both faults and earthquakes, that are formed.

The second major theme is the central role of the rate-state friction laws in earthquake mechanics. These friction laws are now known to not only produce the earthquake instability itself but to result in a gamut of other earthquake phenomena: seismic coupling and decoupling, pre- and postseismic phenomena, earthquake triggering, and the relative insensitivity of earthquakes to transients such as earth tides. Thus the friction laws provide a unifying strand for understanding the commonality of many phenomena previously thought to be disparate. Meanwhile the physics behind these friction laws has become better understood, rendering them less opaque than previously.

The development and deployment of telemetered networks of broadband digital seismometers and of space-based geodesy with GPS and InSAR has provided far more detailed descriptions of earthquakes and the earthquake cycle than ever before. These observations have allowed for their inversion for the internal kinematics of large earthquakes in well monitored regions like California as well as detailed descriptions of interseismic loading and postseismic relaxation, all of which has improved our understanding of the underlying dynamics.

Many people have helped in my preparation of this revised edition. I am particularly indebted to Masao Nakatani, who offered many comments on shortcomings of the first edition and who helped me to better understand the physical basis of the rate-state-variable friction laws.

Palisades, New York, April, 2001
C. H. S.

Preface to the third edition

It has been almost 30 years since I put the finishing touches on the first edition of this book. Since that time there has been enormous progress made on many aspects of the study of earthquakes and faulting. This has required major revisions to be made of all but the first chapter of this book. It is gratifying that these many developments have provided a deeper and broader understanding of brittle tectonics under the unifying application of the principles of rock mechanics. Many of the gaping holes in our understanding that were obvious in the first edition have now been filled and new phenomena have been discovered.

Extensive study has been made of the friction of lamellar minerals such as phyllosilicates. This work has revealed why and under what conditions friction of such materials may be anomalously low compared to the friction of bulk-structure rocks and minerals. This provides a framework for understanding why there are two classes of faults: those that are weak and creeping, and those that are strong and seismogenic. This work also provides a better understanding of the seismogenic properties of subduction zones.

The widespread implementation of the space-based geodetic technologies CGPS and InSAR have provided unparalleled new observations of all stages of the crustal deformation cycle, as well as the discovery of slow slip events in subduction zones and elsewhere. Seismological imaging of earthquakes in the digital age has provided details of the statics and dynamics of earthquakes in greater detail than ever before. These technical developments have come to fruition within a period of the strongest flurry of great subduction earthquakes since the 1960s. The result has been a re-evaluation of the mechanics of these greatest of all earthquakes. Oceanic earthquakes have also been the subject of renewed attention, both through global teleseismic studies and, with the advent of OBS and hydrophone deployments, close-in studies. Whereas the seismogenic properties of continental faults are largely determined by the frictional properties of quartzo-feldspathic rocks and that of subduction zones by friction of metamorphosed phyllosilicates, the distinctive seismogenic properties of oceanic faults can be understood as consequences of friction of mafic and ultramafic rocks.

The preparation of this book has been greatly assisted by reviews of various chapters by colleagues who are experts in those subject areas. Many thanks to Emily Brodsky, Roland

Bürgmann, Mark Anders, Jeff McGuire, and Giulio Di Toro for those reviews. Many thanks also to Maureen Anders, who has drafted many of the new figures in this edition. I particularly am indebted to my wife, Yoshiko, who has supported me in many ways during the preparation of this book.

New York, January, 2018
C. H. S

Symbols

A listing is given of the most important symbols in alphabetical order, first in the Latin, then in the Greek alphabets. The point of first appearance is given in brackets, which refers to an equation unless otherwise noted. In some cases the same symbol is used for different meanings, and vice versa, as indicated, but the meaning will be clear within the context used. Arbitrary constants and very common usages are not listed.

a	atomic spacing [(1.1)], direct frictional velocity parameter [(2.28)]
$a(H_20)$	chemical activity of water [(1.55)]
a-b	combined frictional velocity parameter [(2.28)]
A_r	real area of contact [(2.1)]
b	steady–state frictional velocity parameter [(2.28)], stress dependence of sub–critical crack velocity [(1.55)], exponent in Gutenberg–Richter relation [see below (4.19)]
B	exponent in earthquake size distribution in moment [(4.19)], Skempton's coefficient [paragraph following (4.23b)]
c	crack length [(1.5)]
C	exponent in fault size distribution [(3.9)]
C_0	uniaxial compressive strength [(1.37)]
d	contact diameter [(2.18)]
d_s	jog offset [Section 3.5.1]
D	sliding displacement [(2.21)], specimen size [(1.51)], fault displacement [see below (3.6)]
D_{max}	maximum fault displacement [(3.5)]
D_{ave}	average fault displacement [Figure 3.12]
D_C	critical slip distance [(2.27)]
E	Young's modulus [(1.2)]
$\underline{E}$	effective modulus [(1.9)]
E^*	activation energy [(1.55)]
E_F	frictional work in earthquake [(4.5)]
E_G, G	fracture energy [(4.5)]
E_R	seismic radiated energy [Section 3.3.1]
$F_{ij}(\theta)$	stress function [(1.19)]

$F_i(\theta)$	displacement function [(1.20)]
F	shear force [(2.2)]
F_{SA}	sea anchor force [(6.16)]
F_{SU}	slab suction force [(6.16)]
G	energy release rate [(1.21)], fracture energy [(3.6)]
G_C	fracture energy [(1.24)]
h	hardness parameter [(2.18)]
k, K	stiffness [(2.27)], aftershock productivity [(4.20)]
K_n, K	stress–intensity factor [(1.19)]
K_C	critical stress–intensity factor [(1.24)]
L	length of fault or earthquake rupture, length of slipping patch [(2.36)]
L_C	nucleation patch length [(2.37)]
M_o	seismic moment [(4.1a,b)]
M	magnitude [(4.21)]
M_W	moment magnitude [(4.2)]
n	stress–corrosion index [(1.52)]
N	normal force [(2.1)]
$N(L)$	size distribution of lengths [(3.9)]
N_A	Avogadro's number [(2.32)]
p, p_p	pore pressure [(1.46)], [(6.26)], exponent in Omori law [(4.20)]
$\boldsymbol{p}$	penetration hardness [(2.1)]
P_S	seismic flux [(6.9)]
P_T	tectonic flux [(6.11)]
P_G	seismic flux accumulation [(6.12)]
Q	activation energy [(2.32)]
R	gas constant [(2.32)]
$\boldsymbol{s}$	shear strength [(2.2)]
s	breakdown zone length [Figure 3.10]
t_h	healing time [Section 4.2.2]
t_r	rise time [(2.41)]
T	temperature [(1.55)], thickness of gouge layer [(2.22)], earthquake recurrence time [Section 5.2.2]
T_0	uniaxial tensile strength [(1.39)]
T_1,T_2,T_3,T_4	transition temperatures in crustal strength model [Section 3.4.1]
u, u_i	displacements [(1.20)]
$\Delta\bar{u}$	mean slip in earthquake [(4.1a)]
Δu_i	slip in earthquake [(4.1b)]
v	subcritical crack tip velocity [(1.52)], load point velocity [(2.32)], particle velocity [(2.40)]
v_r, V_R	rupture velocity [Section 2.3.5, (4.14)]
v_{pl}, v_p	plate velocity [Figure 5.11, (6.11)]
$\Delta V/V$	volumetric strain [Figure 1.17]
U	total energy [(1.6)]
U_e	strain energy [(1.6)]
U_s	surface energy [(1.6)]
V	sliding velocity [(2.28)], volume of wear material [(2.21)]
W	work [(1.6)]
W_F	work of faulting [(4.5)]
V_P, α	p wave velocity [Section 4.2.2]
V_S, β	shear wave velocity [Section 4.2.2]
β	shear wave velocity [Section 4.2.2]
γ	specific surface energy [(1.4)]
Γ	Irwin's energy dissipation factor [(1.27)]

δ	joint closure [(2.7)]
δ_{ij}	kronecker delta [(1.45)]
ε_{ij}	strain [(6.2)]
ν	Poisson's ratio [(1.38)]
η	seismic efficiency [(4,7)], viscosity of plastosphere [Section 5.2.2]
η_R	radiation efficiency [(4.9)]
ζ	scaled energy [(4.8)]
θ	state variable [(2.28)]
θ_s	angle of jog [Section 3.5.1]
κ	wear coefficient [(2.22)]
μ	friction coefficient [(2.3)], shear modulus [(1.28)], coefficient of internal friction [(1.34)]
μ_0	base friction coefficient [(2.28)]
μ_d	dynamic friction coefficient [(2.31a)]
μ_s	static friction coefficient [(2.31b)]
μ_{ss}	steady–state friction coefficient [(2.30)]
ρ	radius of curvature [(1.5)], density [(3.2)]
λ	pore–pressure ratio [(3.2)]
σ,σ_{ij}	stress [(1.1)]
σ_t	theoretical strength [(1.1)]
σ_f	Griffith strength [(1.12)]
σ_c	contact normal stress [(2.33a)], critical normal stress [(2.35)]
σ_1	Initial stress [(4.6)]
σ_2	final stress [(4.6)]
σ_A	apparent stress [(4.10)]
σ_F	frictional stress [see below (4.5)]
σ_n	normal stress [(1.33)]
$\overline{\sigma_{ij}}$	effective stress [(1.45)]
$\Delta\sigma_s$	static stress drop [(4.3)]
τ	shear stress [(1.33)], asthenospheric relaxation time [Section 5.2.2]
τ_0	cohesion [(1.34)]
τ_c	contact shear stress [(2.33b)]
χ	seismic coupling coefficient [(6.11)]
Φ	angle of internal friction [(1.35)], dip of subduction interface [(6.16)]
Ω	activation volume [(2.32)]

CHAPTER
ONE

Brittle fracture of rock

Under the low-temperature and pressure conditions of Earth's upper lithosphere, silicate rock responds to large strains by brittle fracture. The mechanism of brittle behavior is by the propagation of cracks, which may occur on all scales. We begin by studying this form of deformation, which is fundamental to the topics that follow.

1.1 THEORETICAL CONCEPTS

1.1.1 Historical

Understanding the basic strength properties of rock has been a practical pursuit since ancient times, both because of the importance of mining and because rock was the principal building material. The crafting of stone tools required an intuitive grasp of crack propagation, and mining, quarrying, and sculpture are trades that require an intimate knowledge of the mechanical properties of rock. The layout and excavation of quarries, for example, is a centuries-old art that relies on the recognition and exploitation of preferred splitting directions in order to maximize efficiency and yield. One of the principal properties of brittle solids is that their strength in tension is much less than their strength in compression. This led, in architecture, to the development of fully compressional structures through the use of arches, domes, and flying buttresses.

Rock was one of the first materials for which strength was studied with scientific scrutiny because of its early importance as an engineering material and in mining. By the end of the nineteenth century the macroscopic phenomenology of rock fracture had been put on a scientific basis. Experimentation had been conducted over a variety of conditions up to moderate confining pressures. The Coulomb criterion and the Mohr circle analysis had been developed and applied to rock fracture with sufficient success that they remain the principal tools used to describe this process for many engineering and geological applications.

 https://doi.org/10.1017/9781316681473.004

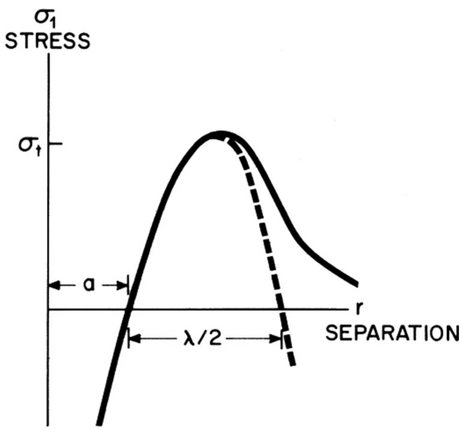

Fig. 1.1. Sketch of an anharmonic model of interatomic forces, showing the relationship between stress and atomic separation (solid curve) and a sinusoidal approximation (dashed curve).

The modern theory of brittle fracture arose as a solution to a crisis in understanding the strength of materials, brought about by the atomic theory of matter. In simplest terms, strength can be viewed as the maximum stress that a material can support under given conditions. Fracture (or flow) must involve the breaking of atomic bonds. An estimate of the *theoretical strength* of a solid is therefore the stress required to break the bonds across a lattice plane.

Consider a simple anharmonic model for the forces between atoms in a solid, as in Figure 1.1, in which an applied tension σ produces an increase in atomic separation r from an equilibrium spacing a (Orowan, 1949). Because we need only consider the prepeak region, we can approximate the stress–displacement relationship with a sinusoid,

$$\sigma = \sigma_t \sin\left[\frac{2\pi(r-a)}{\lambda}\right] \tag{1.1}$$

For small displacements, when $r \approx a$, then

$$\frac{d\sigma}{d(r-a)} = \frac{E}{a} = \frac{2\pi}{\lambda}\sigma_t \cos\left[\frac{2\pi(r-a)}{\lambda}\right] \tag{1.2}$$

but because $(r-a)/\lambda \ll 1$, the cosine is equal to 1, and

$$\sigma_t = \frac{E\lambda}{2\pi a} \tag{1.3}$$

where E is Young's modulus. When $r = 3a/2$, the atoms are midway between two equilibrium positions, so by symmetry, $\sigma = 0$ there and $a \approx \lambda$. The theoretical strength is thus about $E/2\pi$. The work done in separating the planes by $\lambda/2$ is the specific surface energy γ, the energy per unit area required to break the bonds, so

$$2\gamma = \int_0^{\lambda/2} \sigma_t \sin\left[\frac{2\pi(r-a)}{\lambda}\right] d(r-a) = \frac{\lambda\sigma_t}{\pi} \tag{1.4}$$

which, with $\sigma_t \approx E/2\pi$, yields the estimate $\gamma \approx Ea/4\pi^2$.

The value of the theoretical strength from this estimate is 5–10 GPa, several orders of magnitude greater than the strength of real materials. This discrepancy was explained by the postulation and later recognition that all real materials contain defects. Two types of defects are important: cracks, which are surface defects; and dislocations, which are line defects. Both types of defects may propagate in response to an applied stress and produce yielding in the material. This will occur at applied stresses much lower than the theoretical strength, because both mechanisms require that the theoretical strength be achieved only locally within a *stress concentration* deriving from the defect. The two mechanisms result in grossly different macroscopic behavior. When cracks are the active defect, material failure occurs by its separation into parts, often catastrophically: this is brittle behavior. Plastic flow results from dislocation propagation, which produces permanent deformation without destruction of the lattice integrity.

These two processes tend to be mutually inhibiting, but not exclusive, so that the behavior of crystalline solids usually can be classed as brittle or ductile, although mixed behavior, known as semibrittle, may be more prevalent than commonly supposed. Because the lithosphere consists of two parts with markedly different rheological properties, one brittle and the other ductile, it is convenient to introduce two new terms to describe them. These are *schizosphere* (literally, the broken part) for the brittle region, and *plastosphere* (literally, the moldable part) for the ductile region. In this book we will assume, for the most part, that we are dealing with purely brittle processes, so that we will be concerned principally with the behavior of the schizosphere.

1.1.2 Griffith theory

All modern theories of strength recognize, either implicitly or explicitly, that real materials contain imperfections that, because of the stress concentrations they produce within the body, result in failure at much lower stresses than the theoretical strength. A simple example, Figure 1.2(a), is a hole within a plate loaded with a uniform tensile stress σ_∞. It can be shown

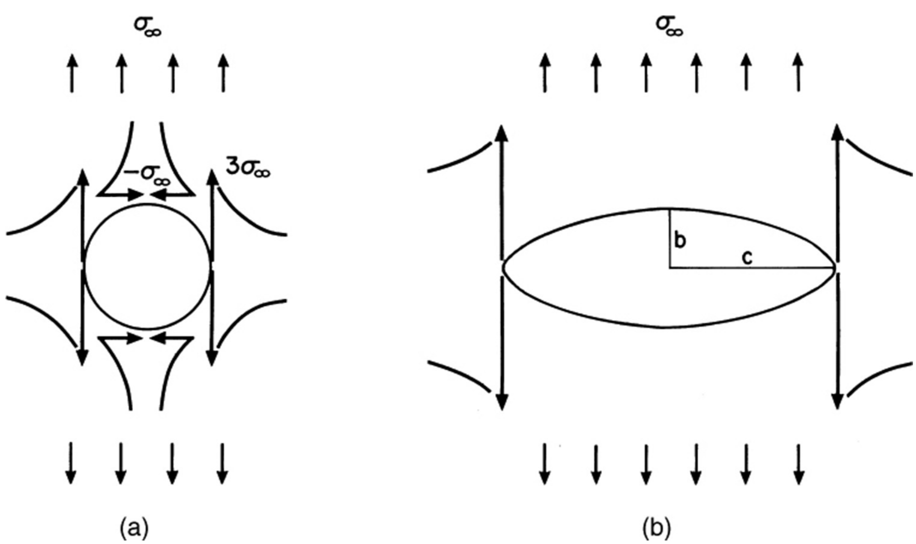

Fig. 1.2. Stress concentration around (a) a circular hole, and (b) an elliptical hole in a plate subjected to a uniform tension σ_∞.

from elasticity theory that at the top and bottom of the hole a compressive stress of magnitude $-\sigma_\infty$ exists and that at its left and right edges there will be tensile stresses of magnitude $3\sigma_\infty$. These stress concentrations arise from the lack of load-bearing capacity of the hole, and their magnitudes are determined solely by the geometry of the hole and not by its size. If the hole is elliptical, as in Figure 1.2(b), with semiaxes b and c, with $c > b$, the stress concentration at the ends of the ellipse increases proportionally to c/b, according to the approximate formula

$$\sigma \approx \sigma_\infty(1 + 2\ c/b)$$

or

$$\sigma \approx \sigma_\infty\left[1 + 2(c/\rho)^{1/2}\right] \approx \sigma_\infty(c/\rho)^{1/2} \tag{1.5}$$

for $c \gg b$, where ρ is the radius of curvature at that point. It is clear that for a long narrow crack the theoretical strength can be attained at the crack tip when $\sigma_\infty \ll \sigma_t$. Because Equation (1.5) indicates that the stress concentration will increase as the crack lengthens, crack growth can lead to a dynamic instability.

Griffith (1920; 1924) posed this problem at a more fundamental level, in the form of an energy balance for crack propagation. The system he considered is shown in Figure 1.3(a) and consists of an elastic body that contains a crack of length 2 c, which is loaded by forces on its external boundary. If the crack extends an increment δc, work W will be done by the external forces and there will be a change in the internal strain energy U_e. There will also be an

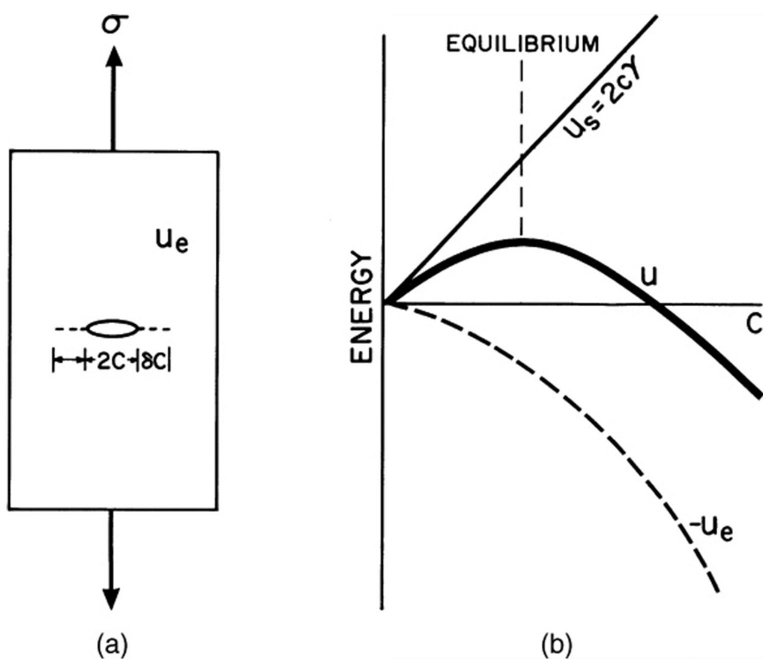

Fig. 1.3. Griffith's model for a crack propagating in a rod (a), and the energy partition for the process (b).

expenditure of energy in creating the new surfaces U_s. Thus, the total energy of the system, U, for a static crack, will be

$$U = (-W + U_e) + U_s \tag{1.6}$$

The combined term in parentheses is referred to as the mechanical energy. It is clear that, if the cohesion between the incremental extension surfaces δc were removed, the crack would accelerate outward to a new lower energy configuration. Thus, mechanical energy must decrease with crack extension. The surface energy, however, will increase with crack extension, because work must be done against the cohesion forces in creating the new surface area. There are two competing influences; for the crack to extend there must be reduction of the total energy of the system, and hence at equilibrium there is a balance between them. The condition for equilibrium is

$$dU/dc = 0 \tag{1.7}$$

Griffith analyzed the case of a rod under uniform tension. A rod of length y, modulus E, and unit cross section loaded under a uniform tension will have strain energy $U_e = y\sigma^2/2E$. If a crack of length $2\ c$ is introduced into the rod, it can be shown that the strain energy will increase an amount $\pi c^2\sigma^2/E$, so that U_e becomes

$$U_e = \sigma^2(y + 2\pi c^2)/2E \tag{1.8}$$

The rod becomes more compliant with the crack, with an effective modulus $\underline{E} = yE/(y + 2\pi c^2)$. The work done in introducing the crack is

$$W = \sigma y(\sigma/\underline{E} - \sigma/E) = 2\pi^2 c^2/E \tag{1.9}$$

and the surface energy change is

$$U_s = 4c\gamma \tag{1.10}$$

Substituting Equations (1.8)–(1.10) into Equation (1.6) gives

$$U = -\pi c^2\sigma^2/E + 4c\gamma \tag{1.11}$$

and applying the condition for equilibrium (Equation (1.7)), we obtain an expression for the critical stress at which a suitably oriented crack will be at equilibrium,

$$\sigma_f = (2E\gamma/\pi c)^{1/2} \tag{1.12}$$

The energies of the system are shown in Figure 1.3(b), from which it can be seen that Equation (1.12) defines a position of unstable equilibrium: when this condition is met the crack will propagate without limit, causing macroscopic failure of the body.

Griffith experimentally tested his theory by measuring the breaking strength of glass rods that had been notched to various depths. He obtained an experimental result with the form of Equation (1.12) from which he was able to extract an estimate of γ. He obtained an independent estimate of γ by measuring the work necessary to pull the rods apart by necking at elevated temperatures. By extrapolating this result to room temperature, he obtained a value that was within reasonable agreement with that derived from the strength tests.

Griffith's result stems strictly from a consideration of thermodynamic equilibrium. Returning to our original argument, we may ask if the theoretical strength is reached at the crack tip when the Griffith condition is met: that is, is the stress actually high enough to break the bonds? This question was posed by Orowan (1949), who considered the stress at the tip of an atomically narrow crack, as described before. Combining Equations (1.3) and (1.4), we obtain

$$\sigma_t = (E\gamma/a)^{1/2} \tag{1.13}$$

This stress will exist at the ends of a crack of length 2 c when the macroscopic applied stress σ_f is (Equation (1.5))

$$\sigma_t = 2\sigma_f(c/a)^{1/2} \tag{1.14}$$

so that

$$\sigma_f = (E\gamma/4c)^{1/2} \tag{1.15}$$

which is very close to Equation (1.12). The close correspondence of these two results demonstrates both necessary and sufficient conditions for crack propagation. Griffith's thermodynamic treatment shows the condition for which the crack is energetically favored to propagate, while Orowan's calculation shows the condition in which the crack-tip stresses are sufficient to break atomic bonds. For a typical value of $\gamma \approx Ea/30$ (Equation (1.4)), commonly observed values of strength of $E/500$ can be explained by the presence of cracks of length $c \approx 1$ μm. Prior to the advent of the electron microscope, the ubiquitous presence of such microscopic cracks was hypothetical, and this status was conferred upon them with the use of the term *Griffith crack.*

Griffith's formulation has an implicit instability as a consequence of the constant stress boundary condition. In contrast, the experiment of Obriemoff (1930) leads to a stable crack configuration. Obriemoff measured the cleavage strength of mica by driving a wedge into a mica book using the configuration shown in Figure 1.4(a). In this experiment the boundary condition is one of constant displacement. Because the wedge can be considered to be rigid, the bending force F undergoes no displacement and the external work done on the system is simply

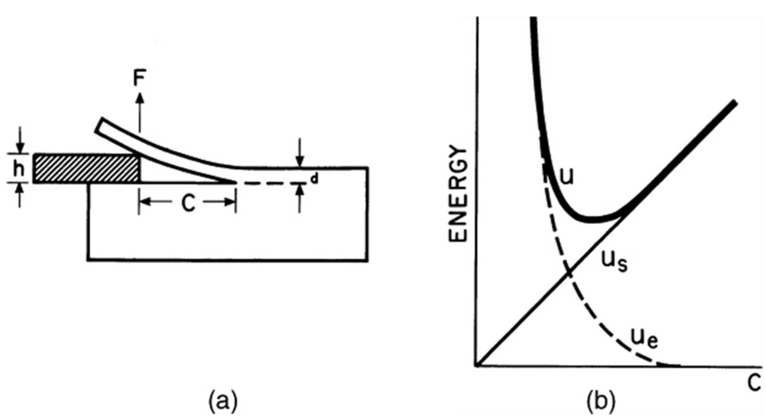

Fig. 1.4. The configuration of Obriemoff's mica cleaving experiment (a), and the energy partition for this process (b).

$$W = 0 \tag{1.16}$$

From elementary beam theory, the strain energy in the bent flake is

$$U_e = Ed^3 h^2 / 8c^3 \tag{1.17}$$

and, using $U_s = 2c\gamma$ and the condition $dU/dc = 0$, we obtain the equilibrium crack length

$$c = \left(3Ed^3 h^2 / 16\gamma\right)^{1/4} \tag{1.18}$$

The energies involved in this system are shown in Figure 1.4(b). It is clear that in this case the crack is in a state of stable equilibrium; it advances the same distance that the wedge is advanced. This example shows that the stability is controlled by the system response, rather than being a material property, a point that will be taken up in greater detail in the discussion of frictional instabilities in Section 2.3. In this case the loading system may be said to be infinitely stiff, and crack growth is controlled and stable. Griffith's experiment, on the other hand, had a system of zero stiffness and the crack was unstable. Most real systems, however, involve loading systems with finite stiffness so that the stability has to be evaluated by balancing the rate at which work is done by the loading system against the energy absorbed by crack propagation.

Obriemoff noticed that the cracks in his experiment did not achieve their equilibrium length instantly, but that on insertion of the wedge they jumped forward and then gradually crept to their final length. When he conducted the experiment in vacuum, however, he did not observe this transient effect. Furthermore, the surface energy that he measured in vacuum was about 10 times the surface energy measured in ambient atmosphere. He was thus the first to observe the important effect of the chemical environment on the weakening of brittle solids and the *subcritical crack growth* that results from this effect. This effect is very important in brittle processes in rock and will be discussed in more detail in Section 1.2.4.

1.1.3 Fracture mechanics

Linear elastic fracture mechanics is an approach that has its roots in the Griffith energy balance, but that lends itself more readily to the solution of general crack problems. It is a continuum mechanics approach in which the crack is idealized as a mathematically flat and narrow slit in a linear elastic medium. It consists of analyzing the stress field around the crack and then formulating a fracture criterion based on certain critical parameters of the stress field. The macroscopic strength is thus related to the intrinsic strength of the material through the relationship between the applied stresses and the crack-tip stresses. Because the crack is treated as residing in a continuum, the details of the deformation and fracturing processes at the crack tip are ignored.

The displacement field of cracks can be categorized into three modes (Figure 1.5). Mode I is the tensile, or opening, mode, in which the crack wall displacements are normal to the crack. There are two shear modes: in-plane shear, Mode II, in which the displacements are in the plane of the crack and normal to the crack edge; and antiplane shear, Mode III, in which the displacements are in the plane of the crack and parallel to the edge. The latter are analogous to edge and screw dislocations, respectively.

If the crack is assumed to be planar and perfectly sharp, with no cohesion between the crack walls, then the near-field approximations to the crack-tip stress and displacement fields may be reduced to the simple analytic expressions: