

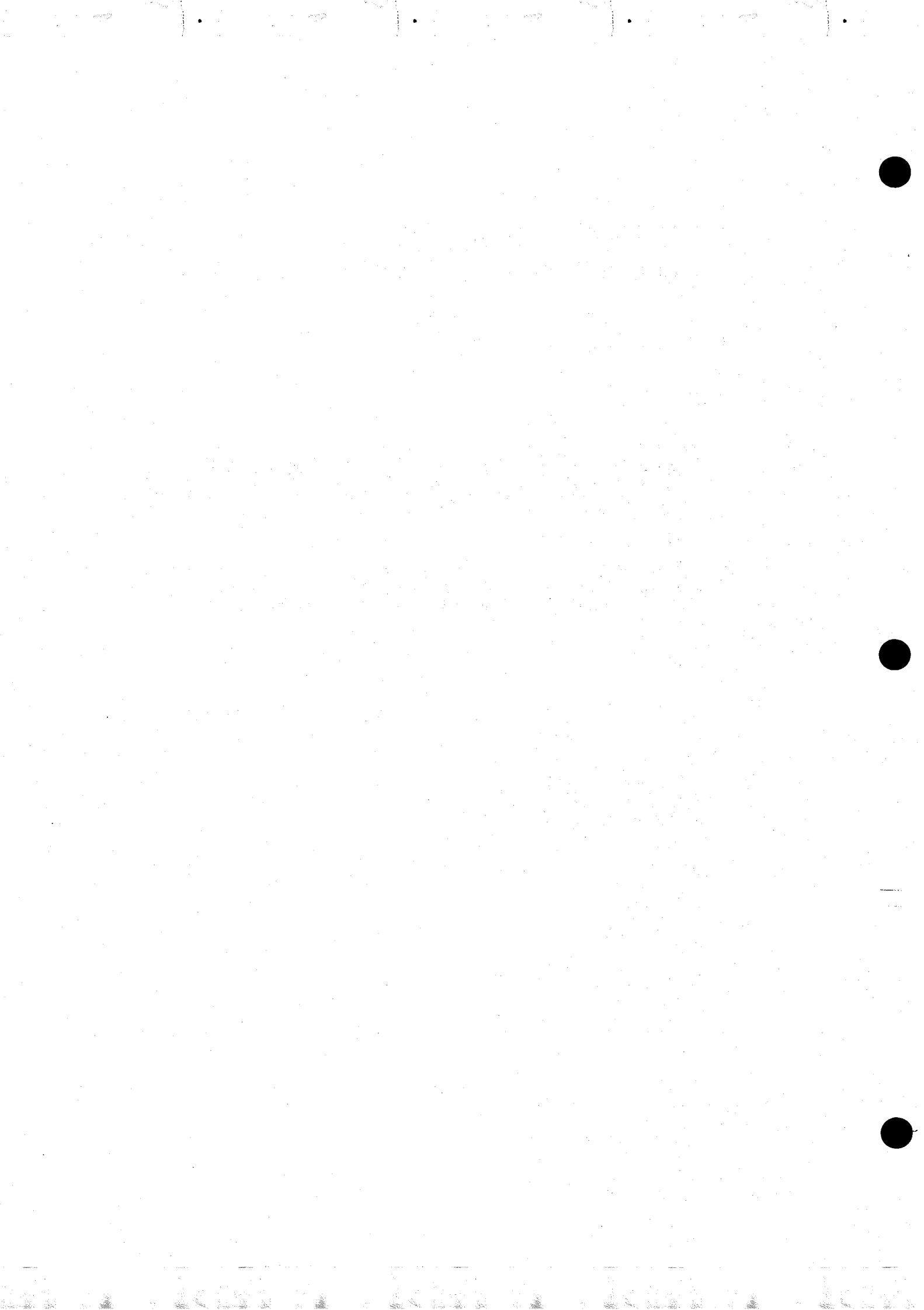


OXFORD MONOGRAPHS ON  
GEOLOGY AND GEOPHYSICS NO. 11

# Impact Cratering

## A Geologic Process

H.J. Melosh



**OXFORD MONOGRAPHS ON GEOLOGY AND GEOPHYSICS NO. 11**

**Series editors**

**H. Charnock  
J. F. Dewey  
S. Conway Morris  
A. Navrotsky  
E. R. Oxburgh  
R. A. Price  
B. J. Skinner**

## OXFORD MONOGRAPHS ON GEOLOGY AND GEOPHYSICS

1. De Verle P. Harris: *Mineral resources appraisal: mineral endowment, resources, and potential supply: concepts, methods, and cases*
2. J. J. Vevers (ed.): *Phanerozoic earth history of Australia*
3. Yang Zunyi, Cheng Yuqi, and Wang Hongzhen: *The geology of China*
4. Lin-gun Liu and William A. Bassett: *Elements, oxides, and silicates: high-pressure phases with implications for the Earth's interior*
5. Antoni Hoffman and Matthew H. Nitecki (eds.): *Problematic fossil taxa*
6. S. Mahmood Naqvi and John J. W. Rogers: *Precambrian geology of India*
7. Chih-Pei Chang and T. N. Krishnamurti (eds.): *Monsoon meteorology*
8. Zvi Ben-Avraham (ed.): *The evolution of the Pacific Ocean margins*
9. Ian McDougall and T. Mark Harrison: *Geochronology and thermochronology by the  $^{40}\text{Ar}/^{39}\text{Ar}$  method*
10. Walter C. Sweet: *The conodonts: morphology, taxonomy, paleoecology, and evolutionary history of a long-extinct animal phylum*
11. H. J. Melosh: *Impact cratering: a geologic process*

# Impact Cratering

## A Geologic Process

H. J. MELOSH

*Lunar and Planetary Laboratory, University of Arizona*

OXFORD UNIVERSITY PRESS • New York  
CLARENDON PRESS • Oxford • 1989

To Ann

*Oxford University Press*

*Oxford New York Toronto  
Delhi Bombay Calcutta Madras Karachi  
Petaling Jaya Singapore Hong Kong Tokyo  
Nairobi Dar es Salaam Cape Town  
Melbourne Auckland*

*and associated companies in  
Berlin Ibadan*

*Copyright © 1989 by Oxford University Press, Inc.*

*Published by Oxford University Press, Inc.,  
200 Madison Avenue, New York, New York 10016*

*Oxford is a registered trademark of Oxford University Press*

*All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means,  
electronic, mechanical, photocopying, recording, or otherwise,  
without the prior permission of Oxford University Press.*

**Library of Congress Cataloging-in-Publication Data**

*Melosh, H. J.*

*Impact cratering: a geologic process / H. J. Melosh.  
p. cm.—(Oxford monographs on geology and geophysics; no. 11)*

*Bibliography: p.*

*Includes index.*

*ISBN 0-19-504284-0 (alk. paper): \$45.00 (est.)  
1. Cratering. 2. Planets—Geology. 3. Impact. I. Title.*

*II. Series.*

*QB603.C7M45 1989  
551.4'4—dc19*

88-5353  
CIP

9 8 7 6 5 4 3 2

*Printed in the United States of America  
on acid-free paper*

## PREFACE

Impact cratering has been recognized as an important geologic process for only the last few decades. As recently as 1950 most astronomers believed that the lunar craters were giant volcanos, and all but a few geologists derided the idea that the earth's surface has been scarred by impact structures kilometers in diameter. A vigorous program of planetary exploration in the Apollo era and continued geologic research on earth has changed these views profoundly. It is now recognized that the cratered landscapes of the moon, Mercury, Mars, and many of the solar system's satellites are sculptured predominantly by repeated impacts of all sizes. More than 100 impact structures with diameters ranging up to 140 km have been identified on earth. More recently, evidence has been found that strongly suggests that the Cretaceous era was ended by the impact of a 10-km diameter comet or asteroid, and the newest and most successful theories of the moon's origin suggest that it formed when a Mars-sized protoplanet struck the protoearth 4.5 billion years ago. The planets themselves are believed to have accreted from dust and debris in the early solar nebula by repeated impacts of smaller objects onto growing planetary embryos, and geochemists are just now beginning to realize that nearly all of the material now residing in planets has been processed through high velocity impacts. In the minds of many geologists, impact cratering has risen from complete obscurity to become one of the "most fundamental" geologic processes. One meteoriticist has even suggested that future historians will accord the recognition of impact cratering an equal importance with the development of plate tectonics.

In spite of these major strides forward in the understanding of impact cratering, many geologists, and even planetary scientists, remain ignorant of vital facts about the cratering process and its resulting craters. Many planetary geologists were surprised and disappointed when Apollo 14 failed to return samples of the moon's mantle from the Fra Mauro formation, even though it was well known among experts on cratering that the impact that created Imbrium basin was unlikely to have excavated material from mantle depths. I have long felt that the reason for this widespread ignorance is the lack of any one place where the major facts about impact cratering have been collected together and presented in an acces-

sible way. Most of our detailed knowledge on impact cratering is spread through a large number of papers and reports that appeared in a wide variety of journals and symposia. Only a few reviews of the entire field are available, and they are not only obsolete, but they also do not go into the individual aspects of cratering in sufficient depth for a professional reader. The most comprehensive source of information on impact cratering in recent years has been the 1977 book *Impact and Explosion Cratering*, which is a collection of papers that were presented at a 1976 meeting in Flagstaff, Arizona. The importance of this book is emphasized by the fact that it is known informally as "the blue bible" among workers on impact cratering. For the noninitiate, however, assimilation of the many papers in this book is both difficult and time-consuming. Furthermore, many advances in the field have been made since 1976.

This book is an attempt to bring together the current knowledge on impact cratering in a single and comprehensive treatise. The field of impact cratering is so large at present, however, that I must restrict the scope to topics that seem to be of major importance to planetary geologists. Thus, I have excluded large amounts of material on microcraters, low velocity impacts, petrology, and geochemistry. I do not treat the geology of individual terrestrial or planetary craters in detail—books including these topics are in progress by other authors. My primary focus is on the impact cratering process, and details of specific terrestrial and extraterrestrial craters are discussed only to the extent that they illuminate that process. I have attempted to give the reader a feel for the essential physics of impact cratering and to establish a theoretical framework that he or she can apply to the particular cratering problem at hand.

My intended audience, however, is not primarily physicists. I have tried to aim the book at geologists and planetary scientists, including both professionals working in other areas who find they need to know something about impact craters, and graduate students who wish to acquaint themselves with this important process. I originally intended, therefore, to minimize the number of equations. I soon found this to be impossible without totally trivializing the subject. Most of the equations, however, are stated as formulas, from which the reader may deduce a quantity of interest on the left-hand side by substituting the

values of some other well-defined quantities on the right-hand side. Although I was unable to maintain this scheme throughout the book, the sections where more mathematical sophistication is required are clearly marked and can be skipped by most readers without losing the essential thread of the exposition.

Because my goal is largely expository, I have tried to keep the number of references in the text to a minimum on the grounds that a text page bristling with citations is both distracting and not very enlightening to someone encountering the material for the first time. This was a difficult decision, as some discoveries or insights may thus become divorced from the name of the worker who first attained them. However, I have tried to provide an entry into the literature at the end of each chapter and the assiduous reader who pursues these selected references should eventually be able to track down the authority behind each discovery. The more important discoveries are, however, explicitly referenced in the text, as are papers that, in my opinion, offer important extensions of the text presentation. The decision of whether to include a direct reference or not is somewhat subjective, and I apologize in advance to any of my colleagues who do not see themselves cited as often as they deserve.

One of the advantages of writing the first monograph on a subject (although it adds considerably to the effort of writing!) is the freedom to organize the material without being influenced by previous prejudices on how it should fit together. I have departed from the usage established in earlier reviews in a number of ways. The first chapter is a brief history of the study of impact craters that provides both background for the facts expounded later and a rationale for much of the work that has been performed to date. It is particularly interesting to me that the nineteenth century astronomers' dismissal of the impact origin of lunar craters was well-founded at the time; only an understanding of high-velocity impact mechanics was able to answer their cogent objections. It was a similar lack of appreciation of cratering mechanics that in 1896 led the eminent geologist G. K. Gilbert to reject impact as the process that created Meteor Crater, Arizona. The second chapter is a brief tour of impact craters at size scales ascending from a few microns to a few thousand kilometers. This chapter is meant to acquaint the complete novice with the general appearance of impact craters and provides a conceptual framework for organizing crater morphology as a function of size. The third chapter is an elementary introduction to stress wave physics with an emphasis on those processes important for impact cratering. Most expositions of shock wave physics begin with the Hugoniot equations and derive the properties of strong stress waves from them. In this book I have chosen to acquaint the

geologically oriented reader more gradually, starting from weak pressure and elastic waves, with which I assume most geologists are reasonably familiar, to describing progressively stronger plastic waves, and finally shock waves. The Hugoniot equations thus appear in the middle of the chapter. Readers with physics backgrounds may wish to skip this chapter, although the material in the last half of the chapter is not part of most physics curricula.

Chapters 4, 5, and 8 are the core of the book. These chapters each deal with one of the three main stages recognized in cratering mechanics: contact and compression, during which the meteorite first contacts the target planet (Chapter 4); crater excavation, when the crater itself opens (Chapter 5); and crater modification, when the crater may partially collapse (Chapter 8). Ejecta deposits are treated immediately after excavation in Chapter 6, since the crater's internal collapse has little effect on ejecta that has already been launched. Similarly, scaling (Chapter 7) is treated before collapse because it generally applies only to the transient crater formed at the end of the excavation stage. Multiring basins are described in a separate chapter (Chapter 9) to clearly distinguish them from complex crater types, including peak-ring craters. In the past, many authors have grouped peak-ring craters and multiring basins together, but, for reasons spelled out in the text, I believe that they are distinct.

Chapters 4 through 9 encompass the classic application of planetary impact cratering studies to single craters formed on airless bodies. The first half of Chapter 10 goes beyond this traditional area to discuss craters as elements of planetary landscapes and offers an introduction to crater statistics. In the second half of Chapter 10 this statistical background is applied to the growth and maintenance of regoliths on airless bodies. This latter portion requires somewhat more mathematical sophistication than the rest of the book, but I believe that anyone with some background in elementary statistics and calculus can work through it. Chapter 11 then discusses what little is known about impacts on planets with dense atmospheres, and how the cratering process may be modified in the presence of an atmosphere. The final chapter describes the connections between impact cratering and other terrestrial or planetary geologic processes and attempts to show impact cratering's place in geology and planetary science as a whole. Mathematical details of the derivation of the Hugoniot equations are treated in the first appendix, while a second appendix describes equations of state that have been found useful for impact cratering computations, including tables of equation of state parameters for a variety of materials of geologic interest.

The reader should be warned that many details of the impact cratering process are still not fully

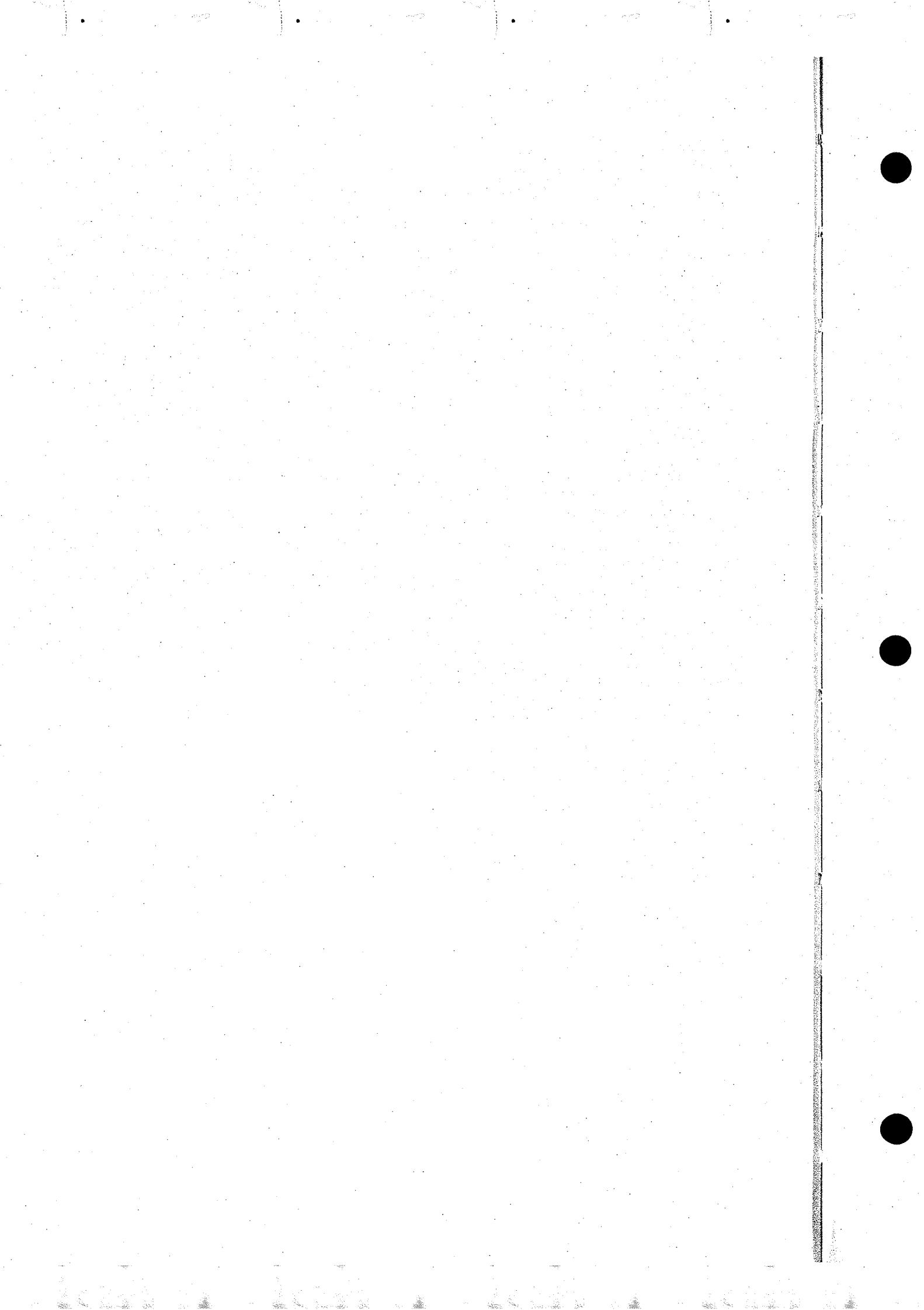
understood. In many cases this leaves room for subjective interpretation. Naturally, I have tended to favor my own interpretations in these cases. Whenever this situation arises in the book, however, I have tried to point out the areas of controversy and have provided references to alternative points of view. Ann Vickery, who read every portion of the manuscript and provided many helpful comments, has expressed the opinion that every expert on impact cratering will find *something* to disagree with in this book. That is as it should be, and by putting these areas of disagreement "on the table" I hope to stimulate further work to clarify these uncertainties. I have also tried to indicate areas, such as oblique impact, that need further work, again in the hope of stimulating further study.

Researching and writing this book have occupied several years. During this time I have had the benefit of advice and assistance from a large number of people without whom this book would never have been completed. Of all these people, Ann Vickery has provided the most consistent support and sometimes the strongest criticisms. This book would have had a very different form without her aid. Several graduate classes at the Lunar and Planetary Lab have tested the success of my pedagogic schemes. Don Gault kindly consented to read the entire manuscript and provided many helpful criticisms along with a great deal of information. He has allowed me to publish a number of photographs, some of which are from his unpublished work. Tom Ahrens read and improved Chapter 4, Robert Schmidt kept me out of

trouble with his review of Chapter 7, and Bob Strom reviewed Chapters 6 and 10 and corrected some of my errors. I owe thanks to Bill McKinnon for unpublished information on the strength-gravity transition, Gene Shoemaker for unpublished data on the primary meteoroid flux, and Ewen Whitaker for historical information on lunar studies, his marvelous prints of Galileo's original lunar maps (the Frontispiece of this book), and the Lick Observatory photo of the full moon. Ed Gaffney provided references to many of the Tillitson equation of state parameters, and Pete Schultz, Ron Greeley, Arthur Frankel, Dick Pike, and Dave McKay all provided sorely needed photographic prints. Mary Ellen Byers excavated large numbers of original photos from the depths of the LPL plate vault and the image library, and Terry Schemenauer did an outstanding job turning them into camera-ready copy. Most of the line drawings in the book are the work of Sue Angelon, although George Feltovic helped in finishing off the last of them as the deadline approached. NASA's Planetary Geology and Geophysics program has supported my research on impact cratering for the past ten years. Finally, I must thank Joyce Berry of the Oxford University Press, whose encouragement changed my vague ideas about a book on impact cratering into a concrete writing project, and whose patient "reminders" helped push the project into completion.

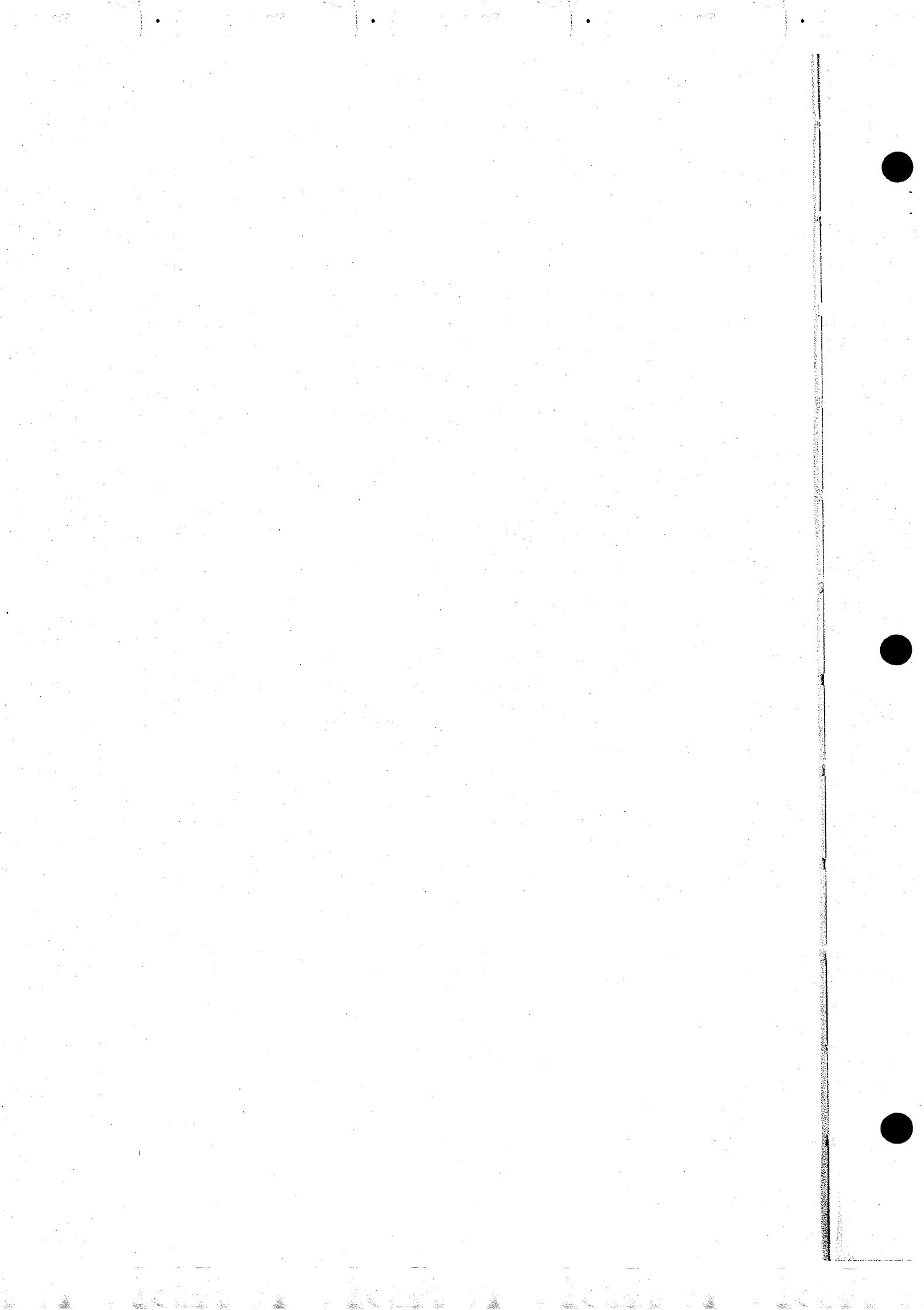
Tucson  
April 1988

H. J. M.



## CONTENTS

1. Themes in the Study of Impact Cratering	3
2. Crater Morphology	14
3. Stress Waves in Solids	29
4. Cratering Mechanics: Contact and Compression Stage	46
5. Cratering Mechanics: Excavation Stage	60
6. Ejecta Deposits	87
7. Scaling of Crater Dimensions	112
8. Cratering Mechanics: Modification Stage	126
9. Multiring Basins	163
10. Cratered Landscapes	184
11. Atmospheric Interactions	205
12. Impact Cratering and Planetary Evolution	215
Appendix I Derivation of the Hugoniot Equations	228
Appendix II Equations of the State for Impact Cratering	230
General Bibliography	235
Index	241



# Impact Cratering

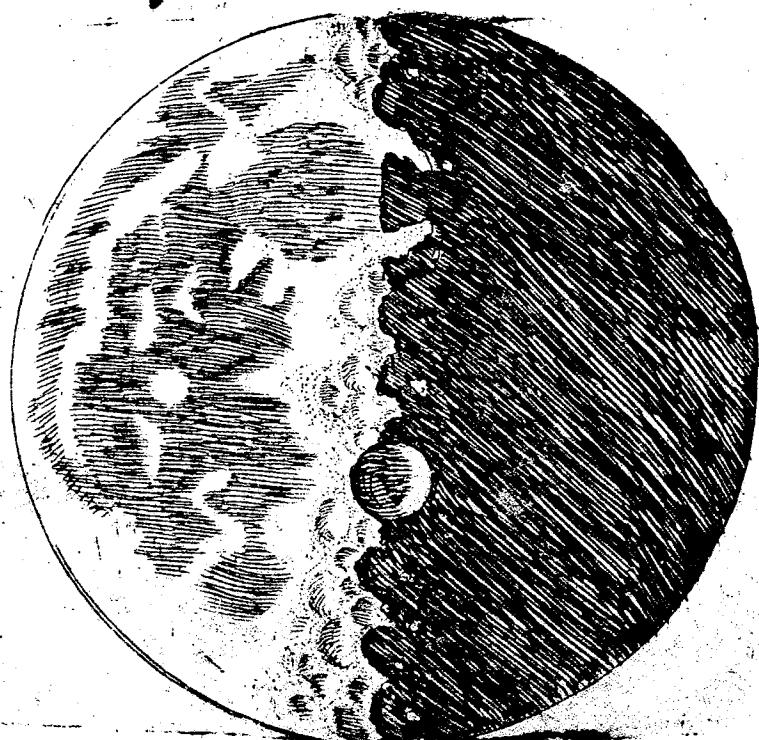
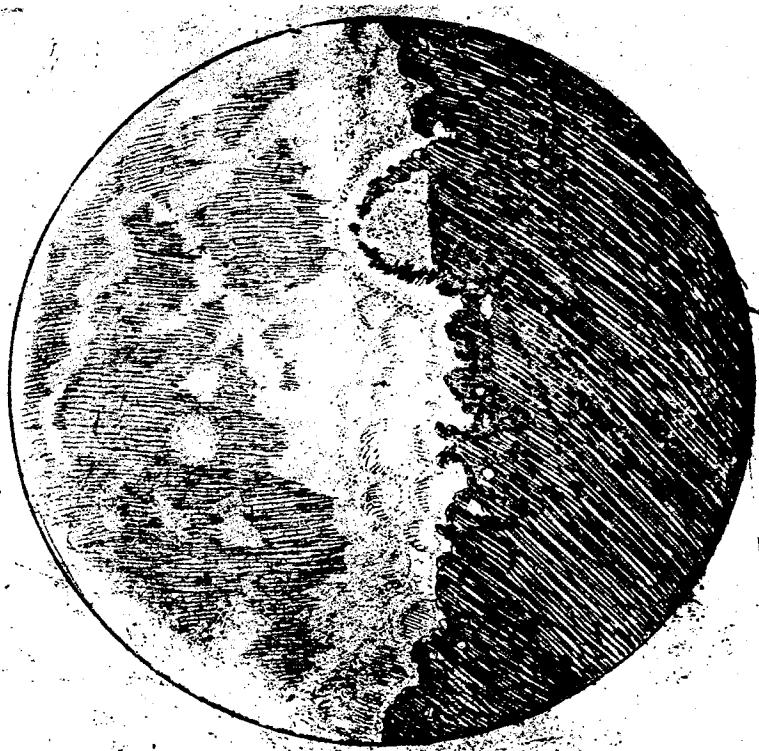
Impact cratering is a geological process that occurs when an object strikes a planetary surface at high velocity, creating a cavity in the target material. This process is responsible for the formation of numerous craters on celestial bodies such as Earth, the Moon, and Mars. The study of impact cratering provides valuable insights into the history and evolution of these worlds.

The impact cratering process can be divided into several stages:

- Entry Phase:** An incoming object, such as a meteorite or comet, enters the atmosphere of a planet or moon. The atmospheric drag causes the object to heat up and glow, creating a fireball.
- Penetration Phase:** The object penetrates the surface of the target body, creating a primary crater. The size of the primary crater depends on the mass and velocity of the impactor.
- Explosion Phase:** The impactor is vaporized or fragmented upon impact, causing a massive explosion. This explosion creates a secondary crater, often larger than the primary one, and sends a plume of debris into the air.
- Refractory Phase:** The impact site is left with a large, irregularly shaped depression. Over time, the crater may be filled with sediment or lava, and the surrounding terrain may undergo erosion and weathering.

Impact cratering has been used as a tool for geological research. By studying the distribution and characteristics of craters on a planetary surface, scientists can determine the age of the surface and the frequency of impacts over time. This information can help us understand the history of the solar system and the evolution of planetary surfaces.

In conclusion, impact cratering is a fundamental geological process that has shaped the surfaces of many celestial bodies. By studying these craters, we gain a better understanding of the history and evolution of our solar system.



Drawings of the moon by Galileo (1610).

# I. THEMES IN THE STUDY OF IMPACT CRATERING

## 1.1 INTRODUCTION

Impact cratering has been recognized as a field of study for only the last few decades. The process itself and its results have, of course, been known far longer. When the history of ideas in this field is examined, it appears that the modern understanding of impact cratering developed out of three initially independent areas of research that joined together at different times. The recent upsurge of interest in impact cratering is largely due to the advent of space travel and, in particular, to the Apollo program.

The oldest and historically best documented of these areas of research is the astronomical study of the origin of lunar craters. Research on lunar craters began with the publication of Galileo's *Sidereus Nuncius* (The Starry Messenger) in 1610. Three centuries passed before the process of impact cratering was properly understood and it took another half century before most of the doubters were convinced that the moon's craters are caused by the impact of large meteorites. Scientific speculation on the craters' origin was not idle during the earlier period, however, and some of the foundations for the remote study of impact craters were laid. The study of extraterrestrial craters continues to this day, except that the principal tool is now spacecraft, not telescopes.

The second area of study began in the early 1880s with the scientific acceptance of meteorites. Small craters produced by meteorites were readily observed at that time, but such craters are produced by projectiles of relatively low velocity and were not regarded as especially important. However, the mere recognition of the existence of meteorites made impact a possible explanation of lunar craters, although few astronomers were willing to defend such a hypothesis. The full connection between meteorites and craters was made in 1906 when D. M. Barringer demonstrated that Meteor Crater, Arizona, is of meteoritic origin. Subsequent geologic work has turned up a substantial number of ancient crater structures on the earth that can be compared to the lunar craters (Grieve, 1987, lists 116 terrestrial impact craters with diameters ranging from tens of meters to 140 km).

The final area of scientific activity that contributed to our modern understanding of cratering processes is both very recent and poorly documented. Explosion craters have long been known

in a military context, but until the beginning of this century they were not studied in any scientific or systematic manner. Military pressures of World War II, however, brought explosion craters under scientific scrutiny. Subsequent concern over the effects of nuclear weapons fueled a large and continuing effort to understand the physics of explosion cratering. High velocity impacts attracted interest in this era because of the anticipated meteoroid hazard to vehicles flying in space.

Study of the physics of impact and explosion craters joined with the astronomical and geological study of impact craters to become the modern field of impact cratering. Work continues in each of these three areas at an undiminishing pace. Indeed, the last few decades have yielded an unprecedented amount of data on extraterrestrial craters, more discoveries than ever of terrestrial impact structures, and better work on the physics of cratering using bigger computers and more sophisticated experiments. There has also been more interdisciplinary work in which experimenters seek to duplicate the forms of extraterrestrial craters, manned landings on the moon have permitted direct geologic study of extraterrestrial craters, and military hydrocodes are being used to compute the outcomes of large impact events.

The field of impact cratering has become so extensive that a complete survey is not possible, even at book length. The complete history of the subject is so large and complex that its exposition would be a full-time undertaking. In the remainder of this chapter I shall merely present the major events in the history of impact cratering and thus put the modern developments in a historic perspective. This work is compiled mainly from secondary sources and I make no pretense that it is complete.

## 1.2 DISCOVERY AND ORIGIN OF LUNAR CRATERS

The study of craters can be said to have begun in 1609 when Galileo Galilei pointed his 1½-in. telescope at the moon and observed circular "spots." Galileo quickly recognized that these spots were depressions because their raised rims were lit before their floors as the terminator swept across them. He also noted the frequent occurrence of central mountains and observed that many of

## IMPACT CRATERING

these depressions were floored with dark material similar to that of the extensive smooth mare. Galileo himself recorded no opinion about the origin of these structures, but later astronomers almost unanimously upheld the idea that they are volcanic.

Galileo's original crude drawings of the moon (Frontispiece) are badly distorted. The large irregular crater in the top part of the figure is probably Serenitatis, and the prominent circular crater in the lower part is Albategnius exaggerated in size for emphasis (for these and other identifications of Galileo's drawings see Whitaker, 1978). It is interesting to note that an Englishman, Thomas Harriot, made a map of the moon about the same time (1609) on which many craters and other features are correctly placed (Kopal and Carder, 1974). However, Harriot failed to publish his map, which has only recently come to light, and so played no role in lunar studies. Better maps than Galileo's were later published by van Langren and Hevelius in 1645 and by Riccoli in 1651.

The first speculations on the origin of the craters were by Robert Hooke in his eclectic book *Micrographia*, published in 1665. Hooke compared lunar craters to pits remaining on the surface of boiled alabaster and suggested that they had formed as a result of internal volcanic activity. Hooke also considered an impact origin, having dropped musket balls and pellets of mud into a mixture of clay and water and noted the similarity of the pits formed by such impacts to lunar craters. However, at that time interplanetary space was thought to be empty, and Hooke dismissed the impact idea because he could not imagine from "whence those bodies should come."

In the years following 1665, lunar studies were largely observational, with the notable exception of the work of J. H. Schröter, who in 1791 published *Selenotopographische Fragmente*, a two-volume treatise that compared lunar and terrestrial topographic features. Schröter measured the heights of lunar features by the lengths of the shadows they cast. After seven years of investigation, Schröter concluded that lunar craters were volcanic, in default of any other known process that produced circular mountain chains. Similarly, Kant in 1785 upheld a volcanic origin, and Herschel in 1787 even reported observing a lunar volcanic eruption.

The next observer after Hooke to suggest an impact origin for lunar craters was Gruithuisen, who in 1829 proposed that the moon's craters had been produced by a cosmic bombardment in past ages. Gruithuisen, however, had previously announced (1822) the discovery of men, animals, and a ruined city on the moon, which he described in minute detail. He was thus not taken seriously and the impact idea waited for further support. It is noteworthy that Gruithuisen's pro-

posal was made only *after* the scientific respectability of meteorites had been established by Chladni in 1819 (see the next section for more details).

The impact idea was almost entirely forgotten after 1837 when N. Beer and J. H. Mädler published an exhaustive study of lunar surface features, including many micrometrically determined crater diameters, peak heights, and an excellent map. Summarizing their work, Neison in 1876 concluded

The moon was to all intents an airless, waterless, lifeless, unchangeable desert, with its surface broken by vast extinct volcanoes.

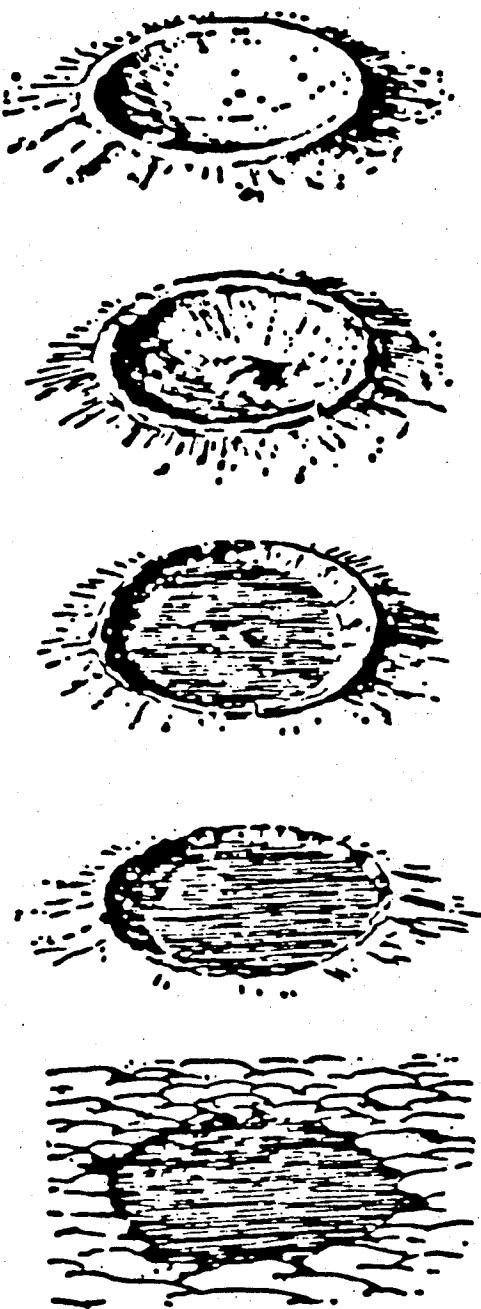
The work of Beer and Mädler, accomplished with a 3½-in. refracting telescope, apparently stifled speculation about the moon by professional astronomers for almost 100 years.

A privy councillor of the mining board in Germany named Althans saw an analogy between lunar craters and artillery craters and in 1839 suggested an impact origin for both lunar and *terrestrial* mountain chains. Naturally, this was not taken seriously.

Geologist J. D. Dana strongly supported the volcanic origin of lunar craters in 1846, after his visit to the Hawaiian islands. Dana offered a new variation of the theory based on his observations of caldera enlargement. Nasmyth and Carpenter in their 1874 book *The Moon: Considered as a Planet, a World, and a Satellite* supported the volcanic hypothesis by detailed comparisons with terrestrial volcanic fields and elaborate drawings of the processes by which lunar craters were supposed to have formed. These works only tended to confirm most astronomers' opinions that the lunar craters are volcanic.

The impact idea was revived again in 1873 by R. A. Proctor in his somewhat popularized book *The Moon*. Proctor's proposal evidently made no impression on the astronomical community and he failed to even mention the idea in the 1878 second edition of the book.

The most serious early support for the impact hypothesis was from G. K. Gilbert (1893) in his study of the moon's face. Gilbert, a famous geologist, described lunar craters and maria from a geologic perspective, recognized a crater size-morphology relation (Fig. 1.1), and tabulated data on depth/diameter ratios. He concluded that only impact could explain the craters. He even suggested that central peaks are due to rebound in a viscous target. Gilbert also recognized the lineations radial to the Imbrium Basin and suggested they are due to scouring by secondary ejecta (although he believed the secondary material to be liquid). He attributed rays to material splashed out by the impact and recognized the wreath of terraces surrounding larger craters as slump features similar to those he had observed on the Col-



**Fig. 1.1** The size-morphology progression recognized by G. K. Gilbert among lunar craters. The top sketch represents the form of the smallest lunar craters, the bottom shows the form of the largest. *From Gilbert (1983).*

orado plateau. (It is interesting that in 1874 Nasmyth and Carpenter had also attributed crater terraces to slumping, although they believed the craters were volcanic in origin.) Gilbert realized that impact at high velocity could produce great heat, computing that an impact at  $1\frac{1}{2}$  miles/second could generate temperatures of  $3500^{\circ}\text{F}$  in the projectile, more than enough to melt it.

Gilbert performed experiments on low velocity

impacts of various powders and slurries in his New York hotel room during the winter of 1891 when he was lecturing on his ideas at Columbia University. The results convinced him that moonlike craters could be formed by impact, but with one proviso—the impact had to be vertical. Many other investigators before and since Gilbert have performed similar experiments. Alfred Wegener, of continental drift fame, wrote a paper on the impact origin of lunar craters in 1920 in which he also reported the results of simple low-velocity impact experiments.

Professional astronomers, however, had a very telling argument against all impact theories, one that was good enough to make impact a very implausible hypothesis at the time. This argument is based on the observation that nearly all lunar craters are circular. Only a tiny fraction of all craters show any perceptible ellipticity. The impact experiments of Gilbert, Wegener, and many others, however, showed that circular craters are produced only by vertical impacts. Oblique impacts invariably produce elliptical craters elongated in the direction of flight. Gilbert tried to get around this problem by postulating that the craters were made by a swarm of objects in orbit around the earth that fell vertically to the moon's surface, but astronomers found this proposal improbable and tended to dismiss the entire impact hypothesis.

Thus, even up to the early years of this century, the impact hypothesis could make little headway. The main reason for this lack of progress was the poor understanding of cratering mechanics. We now know that craters produced by high velocity impacts are circular even at very low angles of incidence, but an appreciation of this fact developed only in the first decades of this century.

E. J. Öpik wrote a prescient paper in 1916 in which he pointed out that the impact of high velocity meteoroids was similar to an explosion and that the craters thus formed would be circular for most angles of incidence. Öpik's paper was published in an Estonian journal, written in Russian with a French summary, and consequently escaped the notice of mainstream astronomers. The obscurity of this publication is usually cited as the reason why astronomers did not appreciate the nature of high speed impact earlier. However, physicist H. E. Ives, who had been working on explosions at Langley Field, Virginia, published a similar idea in the 1919 *Astrophysical Journal*. Ives also realized that a high velocity impact is similar to an explosion and even noted that bomb craters frequently have central peaks that he compared to the central peaks of lunar craters. In Ives' case it was not the obscurity of publication but the daunting criticism of W. W. Campbell, director of the Lick Observatory, that prevented his ideas from being taken seriously.

The true nature of meteoritic impacts was finally impressed upon the scientific world by A. C.

Gifford in two papers published in 1924 and 1930 in the *New Zealand Journal of Science and Technology*. Although this journal appears obscure to American and European scientists, the essence of Gifford's ideas circulated widely in the form of a Hector Observatory *Bulletin*. Gifford also compared impacts to explosions, noting that the kinetic energy per unit mass of a meteorite striking at a speed of 2 miles/second was equivalent to the chemical energy of TNT. Gifford noted that meteoritic impacts should produce circular craters regardless of the angle of incidence. It is interesting that Gifford attributed the apparent absence of meteorite craters on the earth to shielding by the atmosphere. He apparently did not realize that large meteoroids can easily penetrate to the surface.

Subsequent to the 1930s, work on lunar craters was increasingly fused with the geological investigation of terrestrial craters, described more fully in the next section. R. S. Dietz, in 1946, applied geological techniques to the study of the moon's surface features and concluded that lunar craters are of impact origin. The most important event in recent times was the publication of R. B. Baldwin's book *The Face of the Moon* in 1949. Baldwin collected data on the dimensions of lunar, terrestrial, and bomb craters and plotted them together to show the essential similarity of these features, whereas plots of terrestrial volcanic calderas and lunar craters were shown to have little in common. These plots convinced many skeptics that lunar craters are of impact origin. H. C. Urey published a widely read book *The Planets* in 1952 in which he accepted the impact origin of lunar craters. Astronomer G. P. Kuiper became interested in the moon and in 1954 published an influential paper on the moon's surface features that supported the impact origin of craters. Kuiper's work was based on high-quality photographs taken through large telescopes and so impressed many astronomers.

In recent years countless remote investigations of lunar craters at ever-increasing resolution, direct geologic investigation of lunar craters by the Apollo astronauts, and images of craters on planets and satellites throughout the solar system have convinced all but the most hardened skeptics that nearly all the craters on the moon are of impact origin. The discovery of bona fide volcanic calderas on Mars, Venus, and Io has made it clear that volcanoes can be readily recognized on extraterrestrial bodies and differentiated from impact craters.

### 1.3 IMPACT CRATERS ON EARTH

The study of terrestrial impact craters began with the recognition that there are objects in space that occasionally strike the earth at high velocity. Although tales of stones falling from the sky were

well entrenched in popular lore, scientists of the eighteenth century dismissed these as peasant's fables until E. F. Chladni, a physicist best known for his studies of vibrating plates, published memoirs in 1794 and 1819 that suggested that meteors are associated with stony or metallic objects falling from space. Chladni's suggestion was supported by a large meteorite fall on April 26, 1803, when thousands of stones fell near the village of L'Aigle, not far from Paris. Subsequent investigation by members of the French National Institute left little doubt of the reality of the fall. S. D. Poisson suggested that the meteorites were ejected from lunar volcanoes. Asteroids were also discovered about this time, proving that interplanetary space is not as empty as had previously been supposed.

In 1807 at Weston, Connecticut, several meteorites were observed to fall after a brilliant meteor crossed the sky. An apocryphal story promulgated long after his death states that Thomas Jefferson, reflecting the opinion of the informed man of the time, responded to assurances that the fall had been observed by two Yale professors by stating that he "would prefer to believe that two Yankee professors would lie rather than that stones could fall from heaven." (In fact, a case can be made that Jefferson did say "It is all a lie" in reference to the Weston meteorite. The full etymology of this famous quote remains to be told, but a portion of it may be found in Marvin, 1986.) However, many subsequent falls convinced everyone that stones do occasionally fall from heaven. Petrologic and chemical investigations showed that most meteorites are very different from earth rocks, thus supporting an extraterrestrial origin.

The first link between a sizable terrestrial structure and meteorite impact was established by D. M. Barringer in 1906. Barringer convincingly argued that Meteor Crater, Arizona (then commonly known as Coon Mountain), was produced by the high speed impact of a large iron meteorite. Meteor Crater had been studied extensively by G. K. Gilbert in 1891. Although many fragments of meteoritic iron were scattered over the crater rim, Gilbert failed to find any evidence of the large iron mass he expected to lie beneath the crater floor (in this investigation Gilbert used a crude but ingenious magnetic prospecting technique that foreshadowed many modern methods). Because there is abundant evidence for volcanic activity in the vicinity of Meteor Crater, Gilbert persuaded himself that the crater must be of volcanic origin. Thus, the poor contemporary understanding of impact cratering mechanics again misled a competent investigator. Ironically, in this same year Gilbert also convinced himself that lunar craters are created by impacts.

D. M. Barringer's interest in Meteor Crater began as a mining venture. Impressed by the abundant meteoritic iron in the vicinity of the

crater, Barringer staked a claim and started a company to exploit the supposed large meteoritic mass buried beneath the crater floor. Several shafts were drilled and many geologic studies of the crater were completed before Barringer and his colleague B. C. Tilghman published their findings in two separate papers in 1906. Barringer had been impressed by deposits of pulverized rock on the crater rim, which he believed could be produced only by the shock of a meteorite impact, not by volcanism. Tilghman discussed the mechanics of meteoric impact in more detail, and although he correctly estimated the velocity of the impact (between 9 and 42 miles/second), and deduced that the atmosphere would not much hinder a large projectile, he did not appreciate the high specific energy possessed by such a fast projectile and also supposed that a large iron mass must underlie the crater. He dismissed Gilbert's failure to observe a magnetic anomaly over the crater on the grounds that a dispersed mass of iron (he thought) should have a smaller magnetic effect than a homogeneous mass.

Barringer never did find the large mass of meteoritic iron he postulated, even after heroic efforts to sink ever larger and deeper shafts to locate it. As an appreciation of the mechanics of high-velocity impact grew, the mass of iron thought necessary to produce the crater shrank. In 1929 Barringer requested the renowned astronomer F. R. Moulton to perform an analysis of the amount of iron left in Meteor Crater. Barringer at that time was seeking to interest potential financial backers. Moulton independently discovered the impact-explosion analogy and correctly determined that the crater could have been formed by a far smaller projectile than Barringer had supposed (Barringer estimated a mass of 20 million tons, whereas Moulton put the mass between 3 million and 50,000 tons, with a preferred value of 300,000 tons). Moreover, Moulton realized that most of the projectile would be vaporized, not implanted beneath the crater. He also discovered that air resistance is negligible for so large a mass. Moulton's analysis led to a furious but private debate with Barringer and his supporters that was climaxed by Barringer's death from heart failure. It is truly unfortunate that Moulton's results were never published. Only recently has this interesting episode in the history of impact cratering been told in detail (Nininger, 1956; Hoyt, 1987).

In spite of his overestimate of the size of the meteorite that had formed Meteor Crater, at the time of his death in 1929 Barringer had convinced many geologists that meteorite impacts do produce craters on earth. The last serious holdout against the meteoritic origin of Meteor Crater was geologist Dorsey Hager (1953), who published a paper that described the crater as a collapse depression and dismissed the meteoritic iron found in and around the crater as a "coinci-

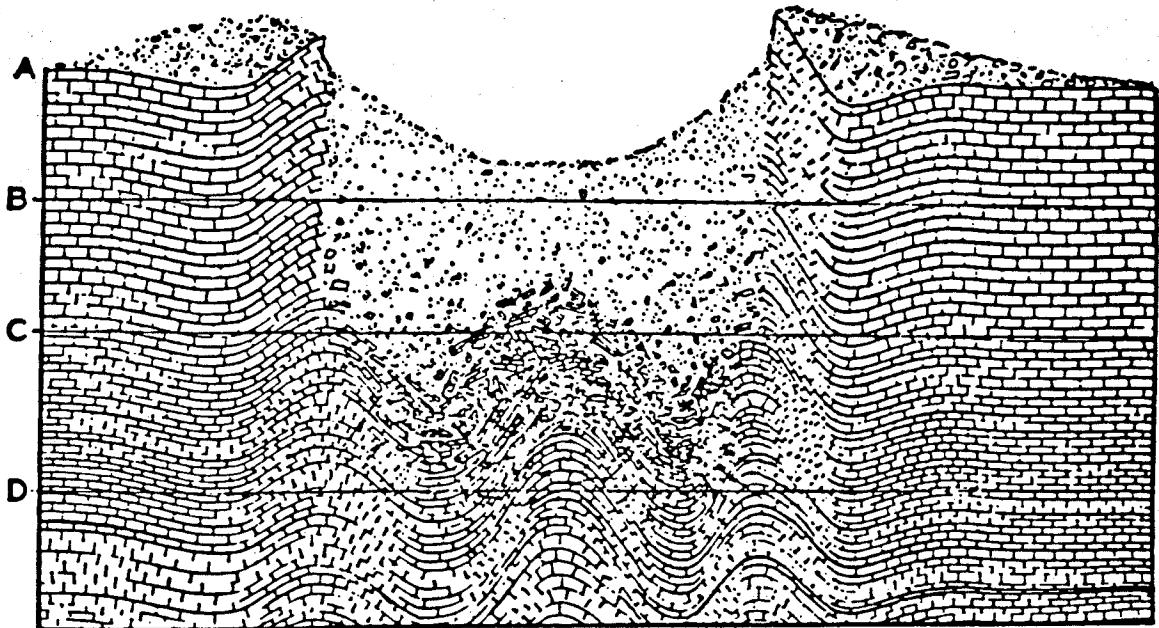
dence." However, it was only a matter of time before more impact craters were found.

Small craters associated with iron meteorites were recognized at Odessa, Texas, by Barringer's son in 1922. The Kaalijarv Crater in Estonia was described by I. A. Reinvaldt in 1927, who ascribed it to an explosion caused by the impact of a high velocity meteorite. Also in 1927 L. Kulik, curator of the meteorite department of the Mineralogical Museum of Leningrad, led an expedition to the site of the great 1908 explosion near the Tunguska River (which had not been visited previously because of its remoteness and the turmoil attendant upon World War I and the Russian Revolution). There he found a devastated forest and microscopic nickel-iron fragments, but no impact craters. The Henbury craters in central Australia were recognized as meteoritic by A. R. Alderman in 1931, again because of their association with meteoritic iron. The discovery of other relatively small craters followed rapidly: Ashanti crater, which contains the circular Lake Botsumtwi in Ghana, was described as an impact crater in 1931 and Wabar crater in Arabia, with its associated iron meteorites, was recognized as meteoritic in 1932. In 1933 L. J. Spencer suggested that the Campo del Cielo craters in Argentina were probably meteoritic (this was confirmed only in 1965).

Larger impact craters were first described as something quite different. The first of these was the 6-km diameter Serpent Mound structure in Ohio, described by W. H. Bucher in 1920. Bucher recognized Serpent Mound as anomalous because of its central area of steeply tilted and uplifted rocks in a region characterized by flat-lying sedimentary rocks. He found the central region had been stratigraphically uplifted some 300 m and was surrounded by a ring-shaped depression. Bucher concluded that the Serpent Mound structure had been the site of some sort of localized explosion. He could not imagine any cause for an explosion other than a volcanic eruption, but was unable to find any volcanic material. In 1924 Bucher visited Steinheim Basin in Germany and recognized the similarity between Steinheim and Serpent Mound. Steinheim had been described as "cryptovolcanic" by W. Branca and E. Fraas in 1905, and the apparently apt name stuck. By 1936 Bucher's interpretation was widely accepted and many more cryptovolcanic structures had been identified. In addition to Serpent Mound and Steinheim, Bucher had described Jeptha Knob in Kentucky, Upheaval Dome in Utah, Decaturville in Missouri, Wells Creek in Tennessee, and Kentland in Indiana. Other geologists described Flynn Creek in Tennessee, Sierra Madera in Texas, and the more than 100-km diameter Vredefort Dome in South Africa as cryptovolcanic.

An appreciation of the nature of impact cratering was growing at this time, however, and in

## IMPACT CRATERING



**Fig. 1.2** Cross section of a typical meteorite crater deduced theoretically by Boone and Albritton. The horizontal lines and letters indicate possible stages in the erosion of an initially fresh crater. *From Boon and Albritton (1937).*

1936 geologists J. D. Boon and C. C. Albritton made the startling proposal that cryptovolcanic structures are, in fact, the scars of ancient meteorite impacts. Boon and Albritton recognized that a meteorite impact is similar to an explosion and reasoned that the localized shattering and uplift characteristic of cryptovolcanic features would also develop when a high velocity meteorite strikes the earth (Fig. 1.2). This hypothesis neatly explains why volcanic materials are absent from cryptovolcanic structures. It also requires that the deformation in cryptovolcanic structures dies out at depth, a prediction that has subsequently been verified for many of Bucher's structures.

The first reliable geologic criterion for the recognition of impact structures in the absence of meteorites was established by R. S. Dietz as a result of his studies of the Kentland structure in 1947. At Kentland Dietz found peculiar fractures in the rock that caused it to break into striated cones. The apex of the cones generally pointed toward the center of the crater, indicating the source of the shock that fractured the rock. Shatter cones had first been recognized by Branca and Faas at Steinheim in 1905, but Dietz argued that they occur only in impact craters, a proposal supported by an experiment in which E. M. Shoemaker and others (1961) produced shatter cones around small-scale impacts in dolomite. In 1964 Dietz announced the discovery of shatter cones around the fabulously rich Sudbury ore body in Ontario and proposed a meteoritic origin for the structure. He also found shatter cones in the Vredefort Dome.

The discovery of Coesite, a high pressure phase

of quartz, in 1953 and a still higher pressure phase, Stishovite, in 1961, provided new geologic criteria for recognizing the high shock pressures that accompany the impact of high speed meteorites. The transformation pressures of Coesite and Stishovite are so high that they have never been observed to form in any volcanic eruption or by any natural process other than impact. Volcanic "explosions" are merely due to the release of pent-up pressure and seem to be incapable of producing shock features in surrounding rocks.

The natural occurrence of Coesite and Stishovite was first recognized by E. M. Shoemaker and his colleagues in rocks from Meteor Crater. In 1961 Shoemaker and E. C. T. Chao reported finding Coesite at the Ries Kessel in Germany, a 22-km diameter crater. Subsequently, Coesite and Stishovite have been found at many other impact sites. Most recently, the discovery of shocked quartz in the Cretaceous-Tertiary boundary clays lends support to the belief that the Cretaceous era was ended by the impact of a 10-km diameter asteroid or comet.

Many more craters were identified in Canada following the 1954 discovery of the New Quebec crater in Labrador by V. B. Meen. Brent crater in Ontario was discovered in 1951, and subsequent investigation by P. M. Millman and others confirmed its meteoritic origin. By 1956 C. S. Beals, M. J. S. Innes, and J. A. Rottenberg had mounted a major campaign to identify and study craters on the Canadian Shield. This program netted such now-familiar craters as Holleford in Ontario, Deep Bay in Saskatchewan, East and West Clearwater lakes in Quebec, and Manicouagan in Que-

bec. As still more craters were discovered, Canadian work shifted toward detailed investigation of each crater, which yielded a rich harvest of results on impact metamorphism, the nature of impact melts, and the detailed structure of both simple and complex craters. M. R. Dence played a prominent part in this program during the 1960s and 1970s. Most recently R. A. F. Grieve has made important contributions to this work, which still continues.

The geologic investigation of impact craters was united with a more complete understanding of cratering mechanics by E. M. Shoemaker. Shoemaker became interested in Meteor Crater in the mid 1950s in connection with his study of the Hopi Buttes diatremes. Beginning detailed investigation in 1957, he compared Meteor Crater to the 1.2-kiloton Teapot Ess underground nuclear explosion crater. Shoemaker's mapping of Meteor Crater led him to recognize the overturned stratigraphy near the crater rim and formulate a theory of impact cratering that he published in 1960 and 1963. This theory is essentially the modern theory of impact cratering, although some additions and corrections have been made since. In 1962 Shoemaker applied his theory to the lunar crater Copernicus and made a strong case for its origin by impact. Shoemaker remains at the forefront of research on the geologic implications of impact cratering.

The major results of the geologic study of terrestrial craters up to the beginning of the Apollo era are presented in a collection of papers published in 1968 as *Shock Metamorphism of Natural Materials*. This book was a landmark at the time and contains many papers that can still be read with profit.

#### 1.4 THE PHYSICS OF CRATERING

Scientific study of the craters produced by impacts and explosions has developed largely in the decades since World War II. Although men like H. E. Ives, an expert on explosions, and B. C. Tilghman, a mathematician and physicist who was an authority on heavy ordinance, contributed to the interpretation of lunar and terrestrial craters, no comprehensive understanding of the cratering process developed until the pressures of war and the development of nuclear weapons spurred governments to support fundamental research on explosion cratering.

Another factor driving an increased interest in high velocity impact was the imminent prospect of space travel. Before 1960 meteoroids were considered to be the major hazard to space vehicles and studies of high velocity impact were undertaken to evaluate and find ways of mitigating the danger. Fear of meteoroids was intensified by measurements made aboard early satellites that showed an apparent meteoroid flux thousands of

times higher than what is now known to be the actual flux. Although it was later realized that these enormous apparent fluxes were due to faulty instrument design, at the time they provided a powerful incentive for the study of high velocity impact.

This period is difficult to evaluate because relatively little has been published on its history. Much of the work was performed under the cloak of secrecy and is only available now in the form of recently declassified reports. A great many people contributed to different aspects of this research and it is sometimes difficult to tell who influenced whom. The Soviet Union also has major programs on cratering mechanics, only a small portion of which are available to the West. The brief history that follows was pieced together from published sources and a few unclassified reports.

There have been two major themes in postwar work on cratering mechanics. The first, and earliest, dealt with the formation of explosion craters. This work was largely focused on craters produced by nuclear explosions and the need to defend military installations from them, although significant advances were also made by the Plowshare Program, which was an attempt to use nuclear explosives for peaceful purposes. The second major theme was the study of small-scale high velocity impacts. Originally motivated by interest in armor-penetrating weapons, this work soon expanded to include the meteoroid hazard to space vehicles and the protection of ICBMs and satellites from high speed debris. Research in this area continues today under the "ballistic weapons" subcategory of the "Star Wars" (Strategic Defense Initiative) program.

The first cratering experiments with large explosive charges may have been conducted in Britain in the 1940s using TNT bombs. In the United States measurements of the pressures, accelerations, and displacements in the ground beneath high-explosive charges were performed by the National Defense Research Committee at Princeton from 1943 to 1945. In 1947 the Panama Canal Company initiated a series of high-explosive tests to study the vulnerability of the Panama Canal to nuclear attack. An important outcome of these studies was C. W. Lampson's 1950 publication of a scaling law that related the diameter and other dimensions of one crater to those of another crater produced by a different-size charge. For the first time it became possible to predict the size of a crater produced by a given explosive energy by observing the size of craters produced by other, generally smaller, energies.

Lampson's scaling relations were refined and extended in subsequent years by a combination of nuclear and high-explosive cratering events. The most significant nuclear events for the study of cratering mechanics were the 1.2-kiloton Jangle U event, detonated in 1951, and the equal-size Tea-

## IMPACT CRATERING

pot ESS, detonated in 1955. An extensive program to delineate the cratering motions was undertaken as part of the Teapot ESS explosion. This involved emplacement and subsequent excavation of large numbers of colored sand columns arranged around the explosion site. Geologic studies of Teapot ESS by E. M. Shoemaker revealed many similarities to his discoveries at Meteor Crater. One of the most significant outcomes of this program of tests was the publication by M. D. Norden of a theory of crater excavation in 1961 followed by an improved scaling law in 1962.

A new direction in the investigation of craters began in the wake of the 1963 Nuclear Test Ban Treaty that prohibited testing of nuclear weapons in the ocean, atmosphere, or space and thus halted any further nuclear cratering explosions by the United States and Soviet Union (France has continued such testing). By this time computers had achieved the capacity to solve the hydrodynamic equations that govern cratering, making it feasible to continue the study of cratering using numerical hydrocodes supported by large high-explosive simulations of nuclear events. The first numerical simulation of a nuclear explosion was performed in 1960 by H. L. Brode and R. L. Bjork of the Rand Corporation. Shortly afterward Bjork performed a numerical simulation of the Meteor Crater impact, published in 1961 (Fig. 1.3). A tra-

dition of numerical computation thus began that has continued to expand to the present day, with ever better computers and codes capable of providing ever more accurate simulations of the cratering process. The results achieved by these methods are epitomized in many of the papers in the important 1977 book *Impact and Explosion Cratering*.

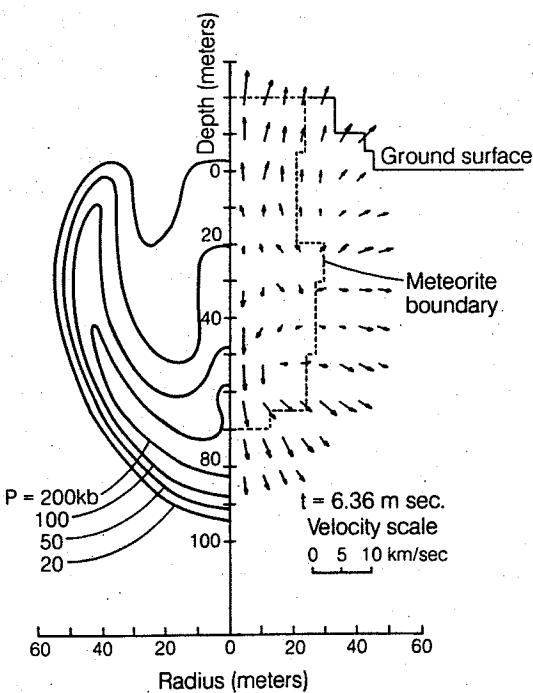
One of the disadvantages of numerical methods is that the results are often difficult to interpret. The sheer mass of data resulting from such computations is difficult for a human to assimilate, and computers are incapable of making generalizations from the specific events computed. An important advance in the study of impact and explosion craters was thus the formulation of the Z-model by D. E. Maxwell in 1973. Maxwell's model was based on the analysis of a large number of hydrocode explosion calculations. It provides an analytic description of the excavation flow field that develops behind the initial shock wave and permits first-order estimates of many aspects of crater growth and ejecta formation.

The most recent work on explosion cratering has returned to the direction of scaling. Using a centrifuge to simulate conditions in large explosion craters, R. M. Schmidt applied sophisticated dimensional analysis to a series of small-scale explosion tests in 1977 and developed scaling relations that describe the results of both explosion cratering experiments and numerical computations with a precision and detail not achieved by any previous scaling relations.

Work on high-velocity impacts began during World War II with the development of armor-penetrating weapons. In these devices a metal cone is collapsed by the detonation of a shaped charge, producing a high-velocity jet capable of penetrating thick armor plate. First described by G. Birkhoff and his colleagues in 1948, this work excited interest in the penetration of solids by high-velocity projectiles. In 1953 N. Rostoker used Birkhoff's theory to estimate the depth and volume of meteorite impact craters and applied his results to Meteor Crater. Rostoker compared his results to the theory of Öpik, which had finally appeared in 1936, and found good agreement.

The cratering theories of Birkhoff et al., Rostoker, and Öpik's 1958 theory focused attention on the hydrodynamic aspects of crater excavation, whereas the explosion analogy emphasized the shock aspects. Although the hydrodynamic and shock wave approaches were regarded as different at the time, modern theories of crater excavation make it clear that both shock and hydrodynamic flow play major roles in the total cratering process.

Increasing understanding of the response of materials to strong shocks grew out of research on nuclear weapon design. High pressure equations of state for metals were devised by theoretical



**Fig. 1.3** Contours of pressure and velocity vectors 6.36 milliseconds after the vertical impact of a 12,000-ton cylindrical iron projectile striking at 30 km/second. This was the first numerical computation of an impact event. *From Bjork (1961).*

physicists and tested in shock wave experiments. Such equations of state are needed for hydrocode calculations of both impacts and explosions. By 1956 knowledge of equations of state had reached a level that permitted J. J. Gilvarry and J. E. Hill to publish a paper on meteorite impact in which they used what is now known as the planar impact approximation to estimate pressures and temperatures early in the impact process. Gilvarry and Hill also concluded that the effective center of the "explosion" caused by the impact is located below the surface at a depth about equal to the projectile diameter.

An important experimental study of impact cratering began in the mid-1950s at NASA's Ames research center at Moffett Field, California. Initially directed by A. C. Charters, this program was intended to evaluate meteoroid hazard and study the formation of lunar craters. The early work of this facility was reported by A. C. Charters and J. L. Summers in a paper published in 1959. Working with metal projectiles and targets, these authors recognized a distinct high velocity cratering regime in which the target behaves as a fluid. They formulated a simple "ballistic pendulum" theory of crater excavation in which the crater size is determined by the projectile's mass and velocity and the target's strength. This laboratory has continued to produce important results, especially since the advent of D. E. Gault's Vertical Gun Range in 1965. The Vertical Gun Range (Fig. 1.4), which is still in operation, permits the acceleration of projectiles to speeds near 6 km/second after which they strike targets in a 2.5-m diameter vacuum chamber where cratering processes may be studied with a variety of techniques.

Work at the Ames Vertical Gun Range produced important theoretical as well as experimental results. In 1963 D. E. Gault and E. D. Heitowitz published a paper on energy partition in high speed impacts. The same year D. E. Gault, E. M. Shoemaker, and H. J. Moore used theory and experimental data to determine the amount of material that impacts might eject from the lunar surface. Scaling laws were established and a theory of impact cratering evolved from this work that was summarized in an influential 1968 paper by D. E. Gault, W. L. Quaide, and V. R. Oberbeck. This theory emphasizes the role of the shock wave and rarefactions in excavating the crater. The amount of work accomplished in the 1960s and 1970s by Gault and his colleagues at Ames is truly staggering and touches nearly every part of the field of impact cratering, especially in the area of the interpretation of extraterrestrial craters. Citations to the work of members of this group occur throughout the remainder of this book and attest to their incredible productivity.

Computational studies of impact cratering began with Bjork's 1961 simulation of Meteor

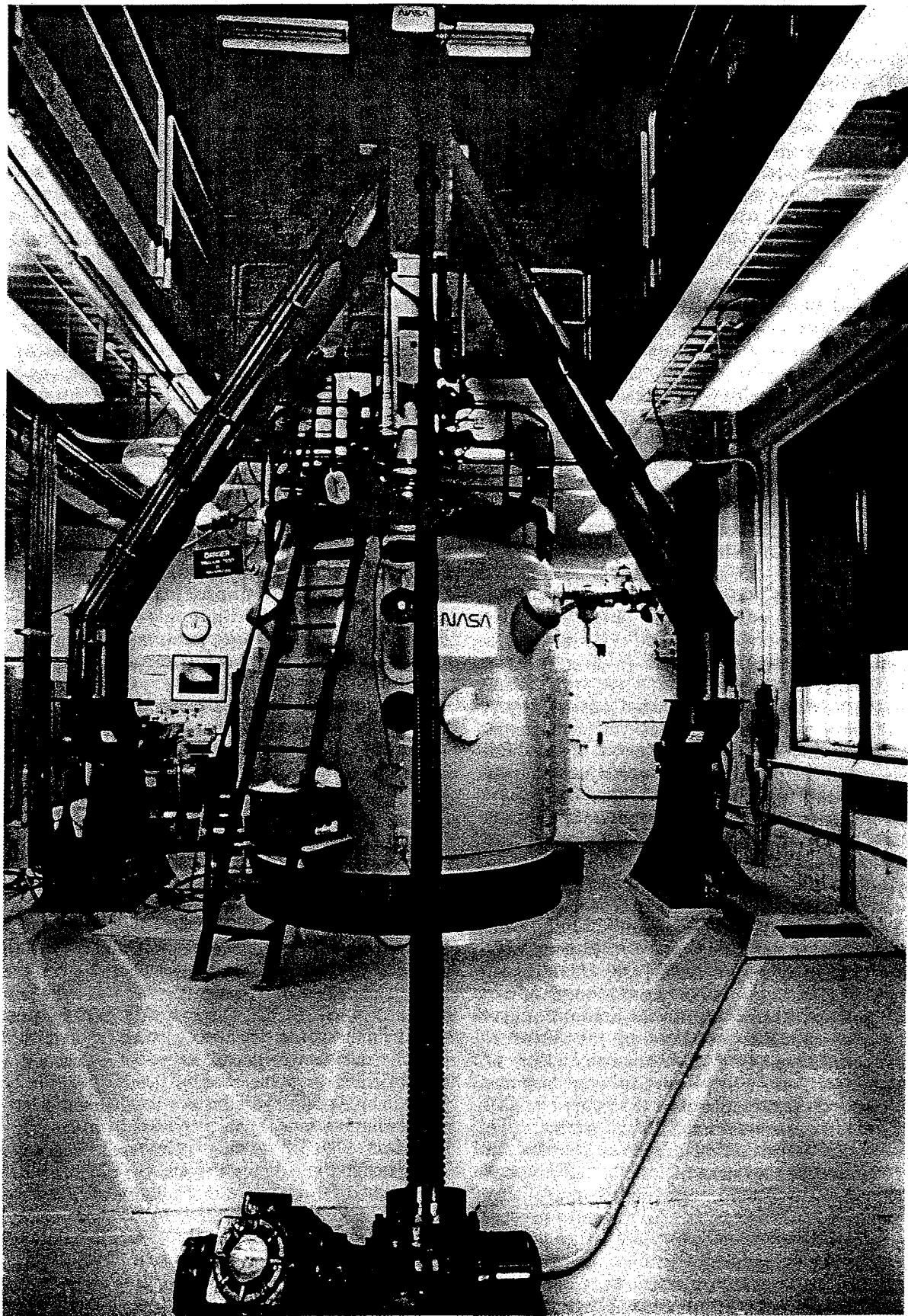
Crater. Eulerian computational methods were developed between 1962 and 1967 by J. K. Dienes and J. M. Walsh, who published a summary of their results in 1970. In this paper they described a scaling relation they called "late stage equivalence," which related the outcomes of impacts at different velocities. Unfortunately, this scaling relation neglected gravity and could not be used for large craters where gravity, not material strength, determines the final crater size. In 1967 R. L. Bjork, K. N. Kreyenhagen, and M. H. Wagner published a numerical study of meteoroid impact for NASA that evaluated the meteoroid hazard. These authors clearly resolved the detached shock, which they differentiated from the excavation flow. Beginning in 1972 T. J. Ahrens and J. D. O'Keefe have continued to publish a series of numerical studies of impact cratering applied to numerous planetary and terrestrial problems. The most recent development in this field is the advent of three-dimensional codes that allow oblique impacts to be computed. Computational work on such impacts is just beginning.

Many other recent developments have occurred in the study of impact craters, such as R. L. Schmidt and K. A. Holsapple's elaboration of late stage equivalence into a comprehensive "coupling parameter" approach to impact scaling laws. However, this work is too new to be properly evaluated in any historical sense. There are many other important pieces of work that have not been mentioned. Such work is the subject of the chapters that follow.

### 1.5 THE IMPACT OF THE APOLLO PROGRAM

The separate themes of lunar crater study, geologic investigation of terrestrial craters, and the physics of cratering were already beginning to intertwine in 1961 when the decision was made to undertake a series of manned landings on the moon. The launch of Sputnik in 1957 and the creation of NASA in 1958 had already increased interest in the cosmic process of impact cratering, but the decision to go to the moon unleashed a flood of effort on the process from all points of view. The Apollo program not only caught scientists' imaginations, it also provided hitherto unheard-of levels of funding to research programs. Facilities like the Ames Vertical Gun Range were established, the U.S. Geological Survey established its Astrogeology branch in Flagstaff, Arizona, under the direction of E. M. Shoemaker, and scientists in university, industry, and government labs were encouraged to do research on problems related to impact cratering.

The unmanned Ranger, Surveyor, and Orbiter missions that reconnoitered the moon prior to manned landings provided a wealth of data on lunar craters at resolutions down to a few milli-



**Fig. 1.4** The Ames Vertical Gun Range at which many experimental studies of impact cratering were performed after the first rounds were fired in January, 1965. The A-frame can carry either a light gas gun or one of several powder guns. The frame can be tilted at different angles to study oblique impacts. The target is placed in the cylindrical vacuum chamber. *Courtesy C. P. Sonett.*

meters. The lunar farside was imaged, first in 1959 by the Soviet spacecraft Luna 3, then later by the U. S. Lunar Orbiter series. Mariner 4 flew by Mars in 1965 and revealed a heavily cratered surface similar to the lunar highlands. The Apollo landings themselves between 1969 and 1972 provided direct samples of the lunar surface. These samples made it clear that impact cratering is a dominant process on the lunar surface and intrudes itself into every aspect of lunar science. The low-altitude orbital flights of the Apollo Command Module provided a platform for remote sensing instruments of many kinds that returned large quantities of information about the lunar surface and, inevitably, about impact craters.

Many of the recent publications on impact cratering can be found in the proceedings of the NASA-sponsored conferences held annually in Houston since 1970. These papers encompass the full range of traditional work on impact cratering in addition to fields such as geochemical studies of lunar crater ejecta that became possible only as a result of the Apollo missions. The recent emphasis in cratering studies has, naturally, turned to studies of craters on planets and satellites that are just coming under study by spacecraft. However, developments at home have not lagged in the wake of the proposal that the Cretaceous era ended with a large impact or that the moon originated from the collision of a Mars-size protoplanet with the protoearth. Impacts on the moon and, probably, Mars have ejected material from the surface of those bodies and supplied us with samples in the form of meteorites.

The Apollo program thus stimulated, accelerated, and supported work on impact cratering. At the present time impact cratering is emerging in many planetary scientist's minds as one of the most fundamental processes in the solar system—a far cry from the time when even the suggestion that the moon's craters were created by impacts was derided as utter lunacy.

## FURTHER READING

### General References

A short historical essay on the origin of lunar craters used to be part of nearly every book on astronomy and the moon. Two books have recently appeared, however, that treat the development of

interest in impact craters in detail. The best of these uses Meteor Crater as the unifying theme:

W. G. Hoyt, *Coon Mountain Controversies*, University of Arizona Press, Tucson, 442 pp. (1987).

The second book, although somewhat popularized, deals more completely with the geologic investigation of impact craters:

K. Mark, *Meteorite Craters*, University of Arizona Press, Tucson, 288 pp. (1987).

The apocryphal Thomas Jefferson quote, as well as other aspects of meteoritics, are dealt with in a delightful paper:

U. B. Marvin, Meteorites, the moon, and the history of geology, *J. Geol. Ed.* 34: 140–165 (1986).

### Special Topics

Alfred Wegner's contribution to the impact theory of lunar craters is not well known to either English or German readers. His original 1920 paper has recently been translated into English and can be found in

A. Wegner and A. M. Celâl Şengör (Trans.), The origin of lunar craters, *The Moon* 14: 211–236 (1975).

An interesting personal view of the development of impact cratering was published by one of recent history's prime movers:

R. B. Baldwin, An overview of impact cratering, *Meteoritics* 13: 364–379 (1978).

Although I am reluctant to cite government reports in this section, the only extensive history of explosion cratering in the United States of which I am aware is

M. D. Nordyke, On cratering: A brief history, analysis, and theory of cratering, *Lawrence Radiation Laboratory Report*. UCRL-6578 (1961).

A brief survey of post-World War II work on high-velocity impacts that emphasizes the military perspective has recently appeared as part of the proceedings of the 1986 Hypervelocity Impact Symposium:

H. Fair, Hypervelocity then and now, *Int. J. Impact Engng.* 5: 1–11 (1987).

A history of numerical code computations focused on impacts is given by

W. E. Johnson and C. E. Anderson, History and application of hydrocodes in hypervelocity impact, *Int. J. Impact Engng.* 5: 423–439 (1987).

## II. CRATER MORPHOLOGY

### 2.1 INTRODUCTION

As early as 1893 G. K. Gilbert had recognized that the detailed morphology of lunar craters is a function of their diameter (see Fig. 1.1). Modern studies with powerful telescopes and spacecraft images have confirmed the existence of a size-morphology sequence on the moon and on other planets and satellites throughout the solar system. Although all impact craters can be described as "circular rimmed depressions," the ancillary details vary not only with size, but also with substrate material, planet, and age.

In this chapter I will briefly describe and illustrate the major features of craters of increasing size. More detailed descriptions of simple and complex craters can be found in Chapter 8. Multiring basins are discussed in Chapter 9. The physics of crater formation is the subject of Chapters 4, 5, and 8. Crater rims and ejecta deposits are treated at length in Chapter 6.

### 2.2 MICROCRATERS

Microcraters were first observed on lunar samples returned to earth by the Apollo astronauts. These tiny pits range from  $0.1\mu\text{m}$  in diameter to sizes large enough to disrupt the rocks on which they are found. Microcraters are evidently due to high velocity impacts of small particles of cosmic dust or, rarely, to small secondary ejecta particles from larger impacts. Figure 2.1 illustrates a typical microcrater. A central bowl-shaped pit is often thickly lined with glass. The pit is frequently surrounded by a light halo of pervasively shattered rock material. On some heavily cratered lunar rocks large numbers of overlapping halos give the outer millimeter of the rock a lighter color than its interior. The pit and halo themselves lie within a much larger depression created by the ejection of spall plates of rock from around the crater. The spall zone is well developed in Fig. 2.1. The removal of material by spalls may be so extensive that the glass-lined central pit appears to sit on a short pedestal.

Craters similar to microcraters have been produced in the laboratory using electrostatic accelerators to generate small, high speed projectiles. Glass appears in the laboratory craters only at impact speeds greater than 10 km/second, from which it is inferred that most microcraters on

lunar samples were formed by fast primary cosmic dust particles. Although most of the focus in this book is on larger craters, the physics of excavation is nearly the same for both microcraters and larger craters. The main difference is that target strength dominates the excavation of microcraters whereas gravity plays the dominant role for most geologically important craters. Another difference is that melt viscosity and surface tension are significant factors in the growth of microcraters. These factors shrink in importance as crater size increases and will thus be neglected in later discussions of cratering mechanics.

Larger craters in the centimeter to meter size ranges are observed in the porous lunar regolith and on large rocks. Craters in very porous materials are typically rimless, irregular pits, whereas meter-scale craters in rock lack the normal breccia fill of larger craters and may exhibit large spall zones. Since such craters are of limited geologic interest (at current photographic resolution on most planets!), I will not discuss them here as a separate class.

### 2.3 SIMPLE CRATERS

The classic type of small crater is the elegant bowl-shaped form known as a "simple crater." Figure 2.2 illustrates the primary characteristics of a simple crater using the 2.5-km diameter lunar crater Linné as an example. This youthful crater is surrounded by a bright ejecta deposit. It has a circular raised rim and an interior slope that is steepest near the rim and smoothly decreases toward the crater's center. There is no central flat floor. The profile of the crater's interior is nearly parabolic. The rim-to-floor depth of such craters is generally about one-fifth of their rim-to-rim diameter, and the rim height is about 4 percent of the diameter.

Simple craters like Linné are created by the impact of a meteoroid moving at speeds exceeding a few kilometers/second. Craters produced by slower-moving projectiles tend to be irregular in plan with broad, less well-defined rims. Many small examples of such craters can be seen in Figure 2.2. These craters are created by secondary ejecta thrown out of large primaries not seen in the photograph. It is usually possible to differentiate secondary craters from primaries on the basis of morphology, unless the secondary has

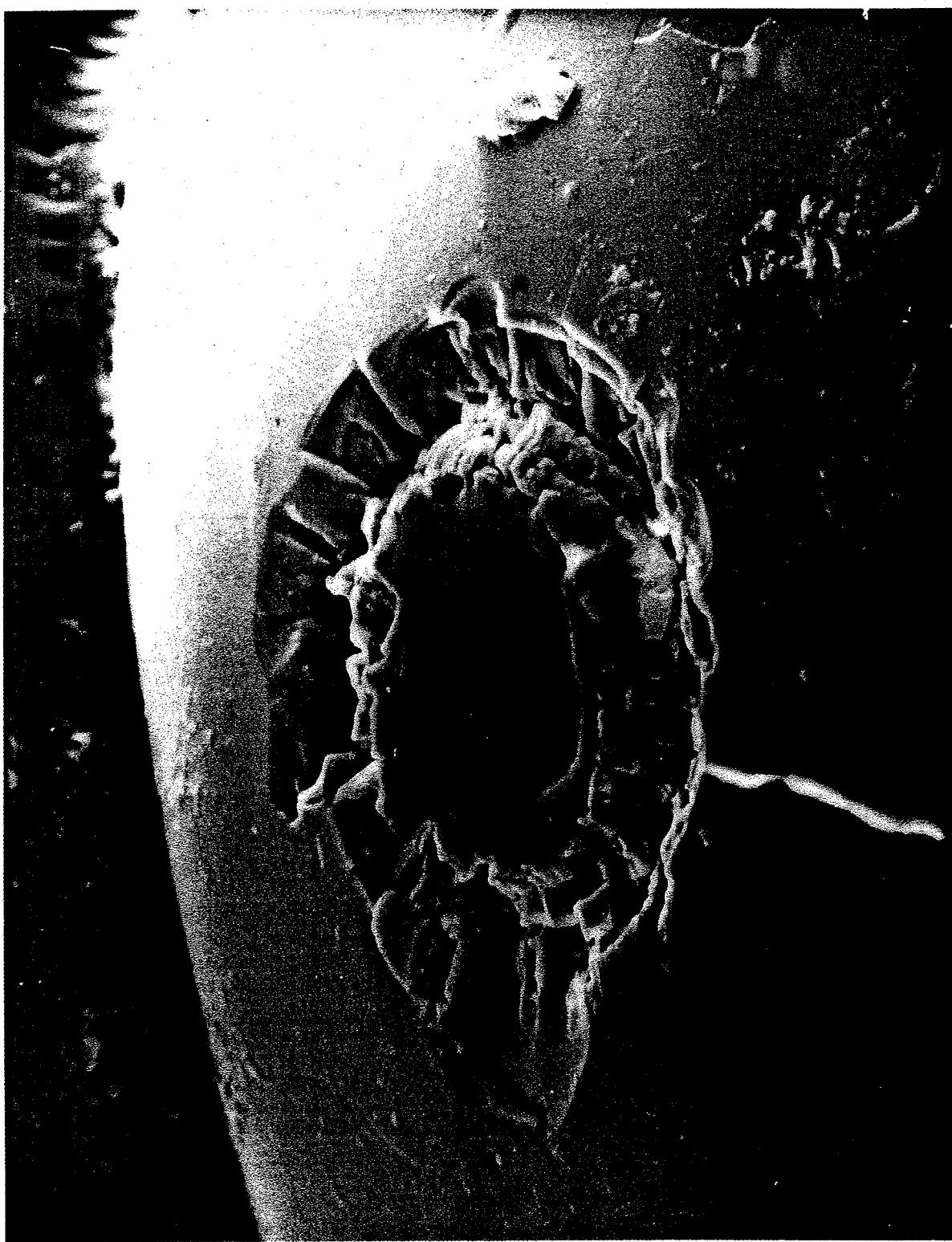


Fig. 2.1 This scanning electron microprobe photograph shows a microcrater  $30 \mu\text{m}$  in diameter on an Apollo 11 glass sphere. The central glass-lined pit is surrounded by an extensive spall zone. NASA S70-18264. (Courtesy D. McKay.)

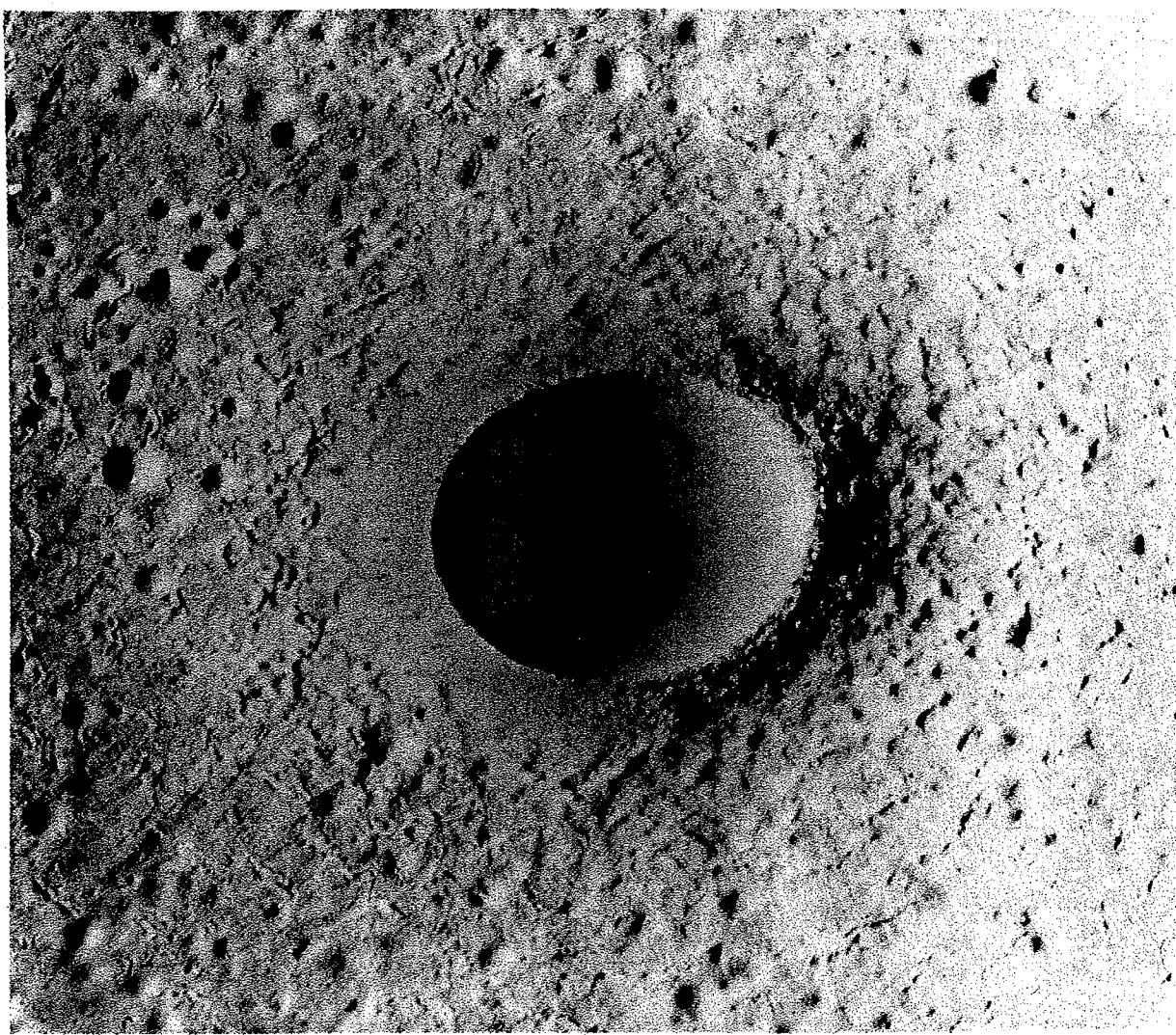


Fig. 2.2 The principal characteristics of a simple crater are illustrated by the 2.5-km diameter crater Linné in western Mare Serenitatis. The bowl-shaped interior is typical of small craters on all planets. *Apollo panoramic photo AS15-9353 (P).*

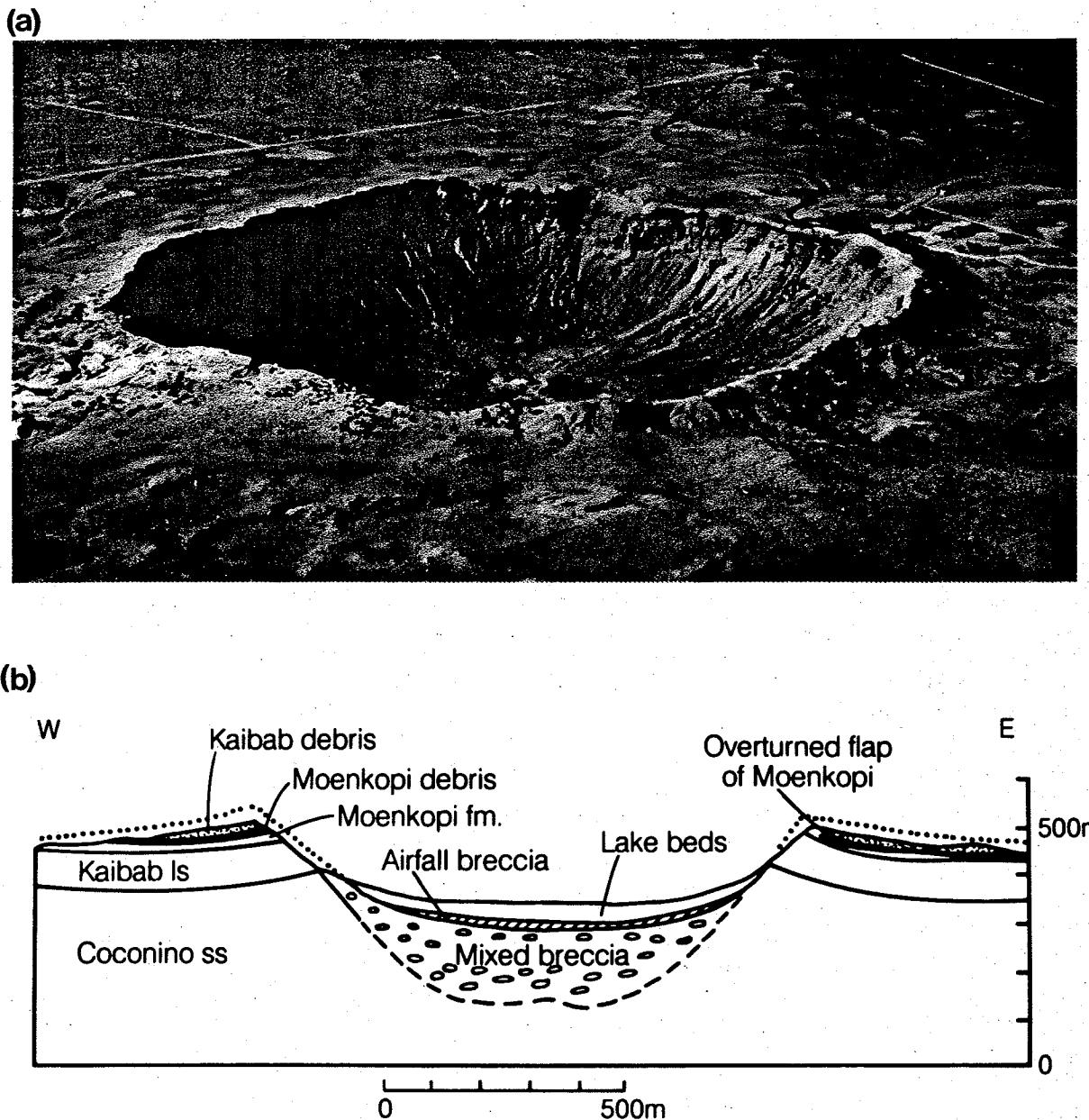
been ejected faster than about a kilometer/second, so that it forms a high-velocity-type crater when it strikes at a great distance from the primary. Secondary craters tend to form in chains and clusters, indicating their common, nearly simultaneous, origin from a larger impact.

Small craters of the simple type are abundant on the moon, Mercury, the satellites of Mars and Jupiter, and all other solar system bodies that have been imaged at high enough resolution. There is no lower limit to the size of simple craters: so long as the target has no significant strength, the simple crater morphology can be observed at arbitrarily small diameters. The upper limit to the size of simple craters is inversely correlated with gravity, and so appears to reflect the onset of some kind of collapse. Simple craters occur on Mars and the earth, but because of their small size are quickly modified by erosion. The

best-exposed and most thoroughly studied simple crater on earth is Meteor Crater in north-central Arizona, United States, on the southern Colorado Plateau.

Meteor Crater (Fig. 2.3a) is thought to have been formed 50,000 years ago by the impact of a 100,000-ton iron meteorite roughly 30 m in diameter which struck at a speed in the vicinity of 20 km/second. The impact site is an area of sedimentary rocks, and the crater may have formed when the water table was well above its present level. There is less impact melt observed at Meteor Crater than at most other impact craters, perhaps because of the presence of water in the porous rocks. There is abundant evidence of high shock pressures, however, from blocks of extensively shock metamorphosed sandstone and limestone.

Figure 2.3b is a simplified geologic cross-sec-



**Fig. 2.3** (a) Meteor Crater, Arizona. (*Roddy and Zeller, U.S.G.S. Courtesy D. Roddy.*) A geologic cross-section of the 1-km diameter crater is shown below in (b). Cross section is after Shoemaker (1960).

tion of Meteor Crater. The crater's rim-to-rim diameter is approximately 1100 m, and its somewhat eroded rim now stands 47 m (originally about 67 m) above the surrounding plain. Its original depth below the plains was about 150 m, but it has been partially filled by Pleistocene lake beds and Holocene slopewash deposits. The crater formed in a flat-lying sequence of Paleozoic rocks consisting of a thin surface veneer or red Moenkopi sandstone underlain by the yellow-white Kaibab limestone, which in turn is underlain by Coconino sandstone. The crater did not penetrate the Coconino sandstone to the deeper-lying Supai formation.

The floor of Meteor Crater is underlain by a lens of mixed breccia some 150 m thick. This breccia is mainly composed of shattered blocks of Coconino sandstone, finely crushed and fused sandstone, and meteoritic material in the form of fine spherules dispersed in glass. Large blocks of Kaibab limestone occur near the top of the lens. These blocks were evidently displaced downward before coming to rest. Deposits of mixed breccia are also found in places on the crater walls, recording an initial episode of upward and outward flow along the crater walls, followed by drainback into the main breccia lens. This lens is overlain by a massive 10-m thick layer of breccia composed of

## IMPACT CRATERING

rock fragments derived from all formations intersected by the crater, along with oxidized meteoritic material. The breccia is graded upward from coarse to fine and is interpreted as an airfall deposit from debris thrown to great height. Approximately 30 m of Pleistocene lake deposits now overlie the airfall breccia. Gravity surveys show a negative Bouguer anomaly centered on the crater, which is presumably associated with the low-density breccia and fractured rock beneath the crater.

The upraised rim of Meteor Crater was created by a combination of two causes. Roughly one quarter to one half of the rim's height is the result of local uplift of the original ground surface around the crater, thus tilting originally horizontal bedding in the walls of the crater outward. This permanent deformation of the rock adjacent to the crater is also seen in nuclear explosion craters and is evidently due to plastic deformation and flow of the rocks surrounding the crater. The remainder of the rim height comes from material ejected from the crater. E. M. Shoemaker first noted in 1960 that ejected material near the rim possesses an inverted stratigraphy. This "overturned flap" of Moenkopi overlain by Kaibab debris, which is in turn overlain by Coconino debris, is illustrated in Figure 2.3b.

Much of Meteor Crater's original ejecta blanket has been removed by erosion. However, even now it is characterized by large blocks of rock on the rim and hummocky topography for several hundred meters beyond it. Fresh lunar craters comparable in size to Meteor Crater have similar blocky rims. Hummocky topography around lunar craters grades into dune forms beyond 1.1 to 1.3 crater radii from the center. This topography extends out to the edge of the continuous ejecta blanket about one crater diameter away from the rim. Secondary craters occur beyond this range (see, e.g., Fig. 2.2).

Small extraterrestrial craters are believed to be similar to Meteor Crater in most respects, except, of course, for the deposits peculiar to the earth's environment, such as the lake beds and airfall breccia. Several small craters investigated geologically by the Apollo astronauts show the basic similarity of all simple craters.

## 2.4 COMPLEX CRATERS

An abrupt change in the interior morphology of lunar craters takes place at a diameter between 10 and 20 km. Figure 2.4 shows that the floor of the 17-km diameter lunar crater Bessel is relatively flat and is covered with what appear to be landslide deposits from the crater walls. The interior walls slope steeply down from the rim, standing close to the angle of repose, and the crater is much shallower for its diameter than a simple crater.

The transition from simple craters is more profound for the 25-km diameter lunar crater Euler

(Fig. 2.5), which exhibits slump terraces on its walls and a central peak in addition to the relatively flat interior floor. It seems evident that these large craters have somehow collapsed compared to smaller simple craters. This notion of collapse is strengthened by the observation that this transition scales inversely as the gravitational acceleration of the planet on which the craters form. Thus, the transition takes place at about 7-km diameter on Mercury, which has twice the moon's surface gravity, and between about 2 and 4 km on the earth, which has six times the moon's gravity.

Craters with wall terraces, central peaks, and flat floors are called *complex craters*. Although the term was originally applied to structurally modified terrestrial craters by M. R. Dence in 1965, it seems that the structures Dence was referring to are equivalent to collapsed lunar craters and the same term should be applied to both. As will be discussed at greater length in Chapter 8, the central peak in craters like Euler is *not* simply debris from landslides that slid down the crater walls and piled up in a central heap. Studies of terrestrial complex craters show that the central peak is composed of rocks that originated below the crater floor and that were uplifted a distance equal to about 8 percent of the crater's final diameter. This large-scale uplift beneath the crater floor marks a profound difference between simple and complex craters.

The characteristics of complex craters are fully developed in the 102-km rim-to-rim diameter lunar crater Theophilus (Fig. 2.6). The flat floor of Theophilus lies 2.8 km below the surrounding plain and its rim rises 1.3 km above it. In planform the rim is scalloped, consisting of a wreath of coalescing alcoves, each about 10 km across. A zone of terraces begins at the base of a high scarp inside the rim, descending stepwise across a zone 15- to 20-km wide to the flat inner floor of the crater. The scarp stands near the angle of repose and the terraces appear to be slump blocks. The widest terrace is located just within the rim and ranges between 5- and 7-km wide. The terrace width decreases inward. Individual terrace segments are typically 10-km long, although their terminations are sometimes obscure. The tops of these terraces are tilted outward and locally trap small ponds of frozen (presumed) impact melt. In the somewhat fresher lunar crater Copernicus (Fig. 6.10) such melt ponds also occur on the rim and leveed channels may be seen descending the walls onto the crater floor. Theophilus's western wall is somewhat modified by its intersection with the older degraded 98-km diameter crater Cyrillus.

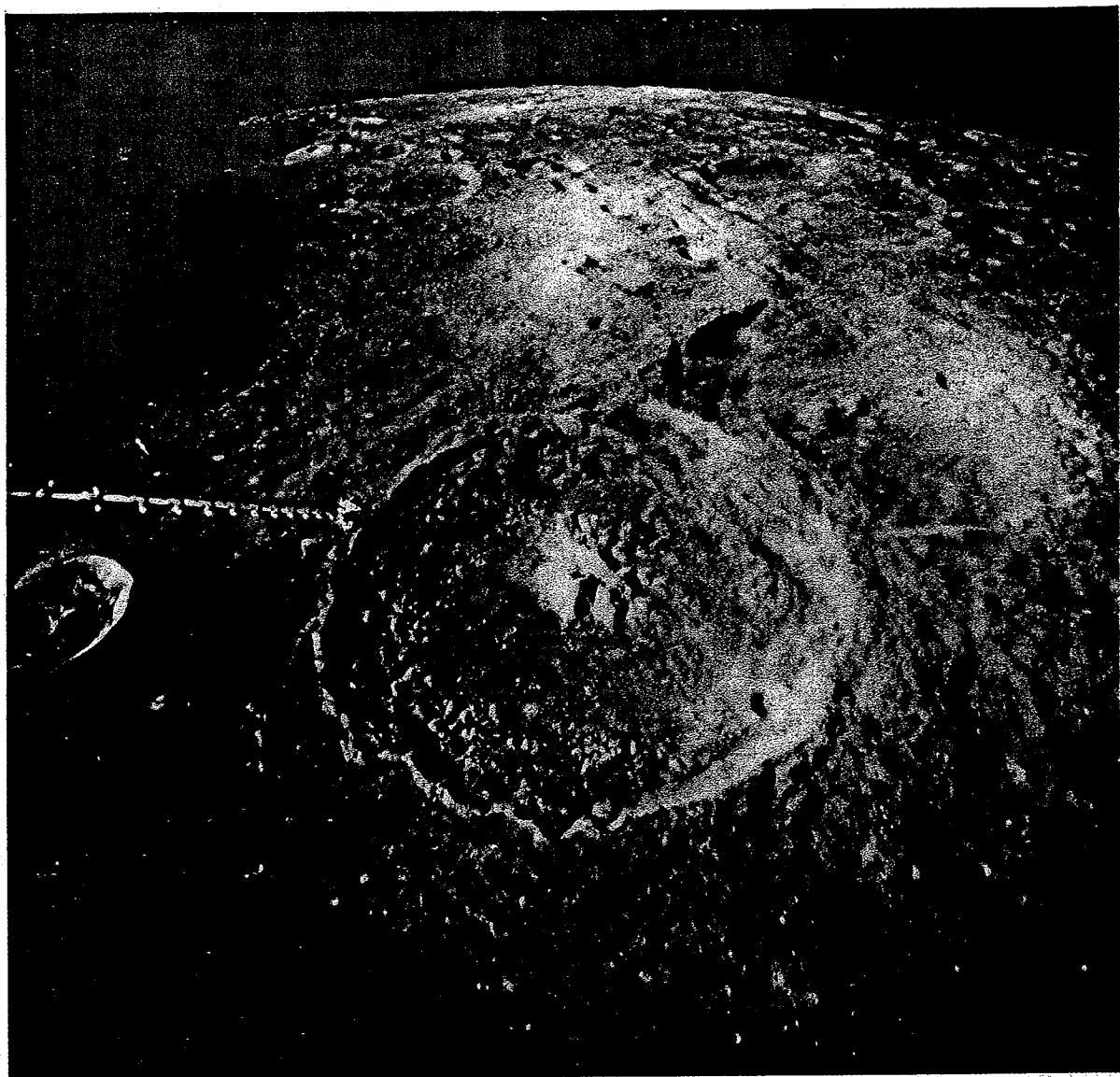
A striking group of central peaks rises nearly 3 km from Theophilus's floor. As in most complex craters the diameter of this central peak complex is about 22 percent of the crater's rim-to-rim diameter. Theophilus's central peaks are unusually



**Fig. 2.4** Crater morphology close to the simple-to-complex transition is exemplified by the 17-km diameter lunar crater Bessel. Note the steep inner wall, on which several rock outcrops may be seen, and the hummocky floor suggestive of massive wall collapse. *Apollo panoramic photo AS15-9328 (P).*



**Fig. 2.5** A central peak and wall terraces appear in the 25-km diameter crater Euler. Rock outcrops appear on its walls and the smooth area on its floor is probably solidified impact melt. Strata beneath terrestrial craters with similar morphology are dramatically uplifted beneath the central peak. *Apollo panoramic photo AS15-0274 (P).*



**Fig. 2.6** Complex crater characteristics are fully developed in the 102-km diameter lunar crater Theophilus. An extensive flat floor veneered with impact melt separates the central mountains from the terraced walls. The development of Theophilus' far wall was influenced by the older degraded crater Cyrilus in the background. *Apollo 16 metric camera frame 0692.*

high: they rise several hundred meters above the level of the surrounding plain, whereas in most lunar craters the central peaks are well below this level. Even in Theophilus, however, the central peaks are lower than the rimcrest.

The central peak morphology is modified in craters somewhat larger than Theophilus. In lunar craters with rim-to-rim diameters of about 140 km and larger the central peak vanishes and is replaced by an inner concentric ring of irregular mountain peaks (see Figures 8.8a and 8.8b). Such peak-ring craters have been recognized on the earth, moon, Mars, and Mercury (Fig. 8.8a and 8.8b), and in all these cases the inner ring diameter is roughly half the rim-to-rim diameter. The

transition diameter, however, scales inversely as gravitational acceleration, in exactly the same way as the simple-to-complex transition diameter scales. The nature of this transition is conjectural, but it is thought to be due to a fluidlike "sloshing" of material inside the crater. The physics of this process is discussed in more detail in Chapter 8.

Peak-ring craters do not form on the icy satellites of Jupiter. Instead, large complex craters on Ganymede and Callisto exhibit central pits (Fig. 8.11) that may or may not be analogous to peak rings. The pits sometimes possess raised rims that probably develop long after crater excavation as a result of isostatic upbowing of the crater floor.

Young complex craters of all sizes and on all

planets have extensive ejecta blankets that typically extend one crater diameter from the crater rim, independent of gravitational acceleration. The distal parts of the ejecta blanket are often radially linedated, showing signs of scour and deposition by rapidly moving debris traveling radially outward from the crater's center. Secondary crater fields are encountered at greater distances from the primary crater. The extent of the secondary crater fields does depend upon gravitational acceleration. Thus, Martian and Mercurian secondary crater fields lie considerably closer to their primaries than lunar secondary crater fields. The causes of these relations are discussed in Chapter 6.

## 2.5 MULTIRINGED BASINS

It is traditional in lunar studies to cite multiringed basins as the next stage in the size-morphology sequence. On the moon this is indeed the case, for large inward-facing scarps such as those surrounding the Orientale Basin (Fig. 2.7) develop at diameters larger than most peak-ring craters. However, study of large impact structures on other planets puts this simple interpretation into doubt. It is unclear whether even the largest impact craters on Mercury (the 1300-km diameter Caloris Basin) and Mars (the 2000-km diameter Hellas) are of the same type as Orientale. Even if some of the larger craters on Mercury and Mars can be shown to be of the lunar multiring type, the transition diameter clearly does not scale as the inverse gravitational acceleration, indicating that the process that creates multiring basins is different from the collapse that induces either the simple-to-complex or peak-to-peak ring transitions.

In this book I am departing from the traditional view that multiring basins are simply a large-diameter development of normal crater collapse processes. This is partly motivated by the discovery of the Valhalla Basin, a distinctly different multiring basin type on Callisto (Fig. 2.8). Both lunar and Valhalla-type multiring basins have been observed on Ganymede. Further discussion of the physics of multiring basin formation is deferred until Chapter 9.

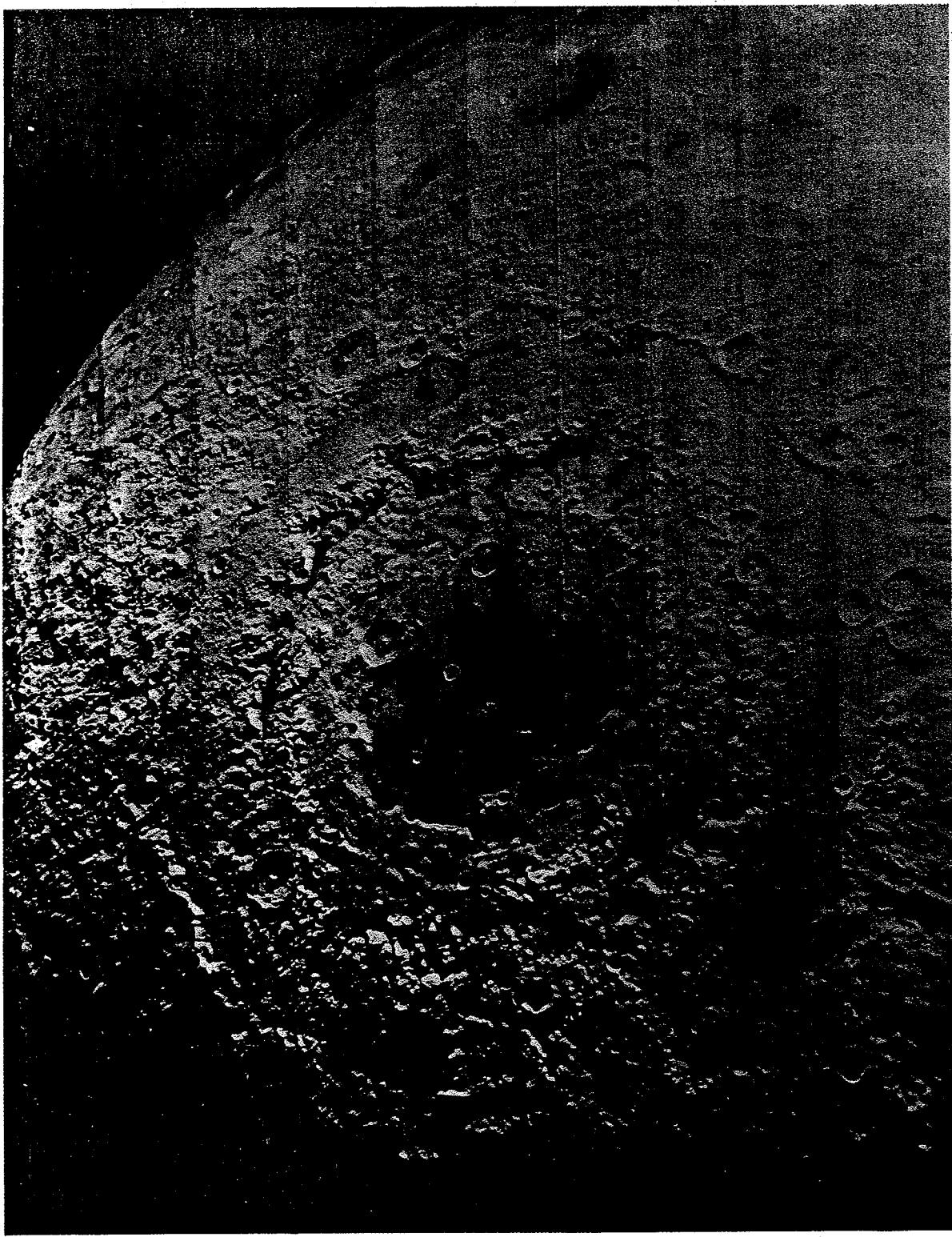
I have also departed from tradition by classifying peak-ring structures as a variety of complex crater. With a few notable exceptions, most previous workers have considered peak-ring craters to be a variety of multiring basin. However, as I have already stated, peak rings *do* develop as part of the normal size-morphology sequence on all of the terrestrial planets. Peak-ring craters seem to arise from a gravity-related extension of the process that raises central peaks in smaller craters, because the ring clearly forms inside the original crater rim in the best-preserved examples, and be-

cause the transition diameter scales inversely as gravitational acceleration.

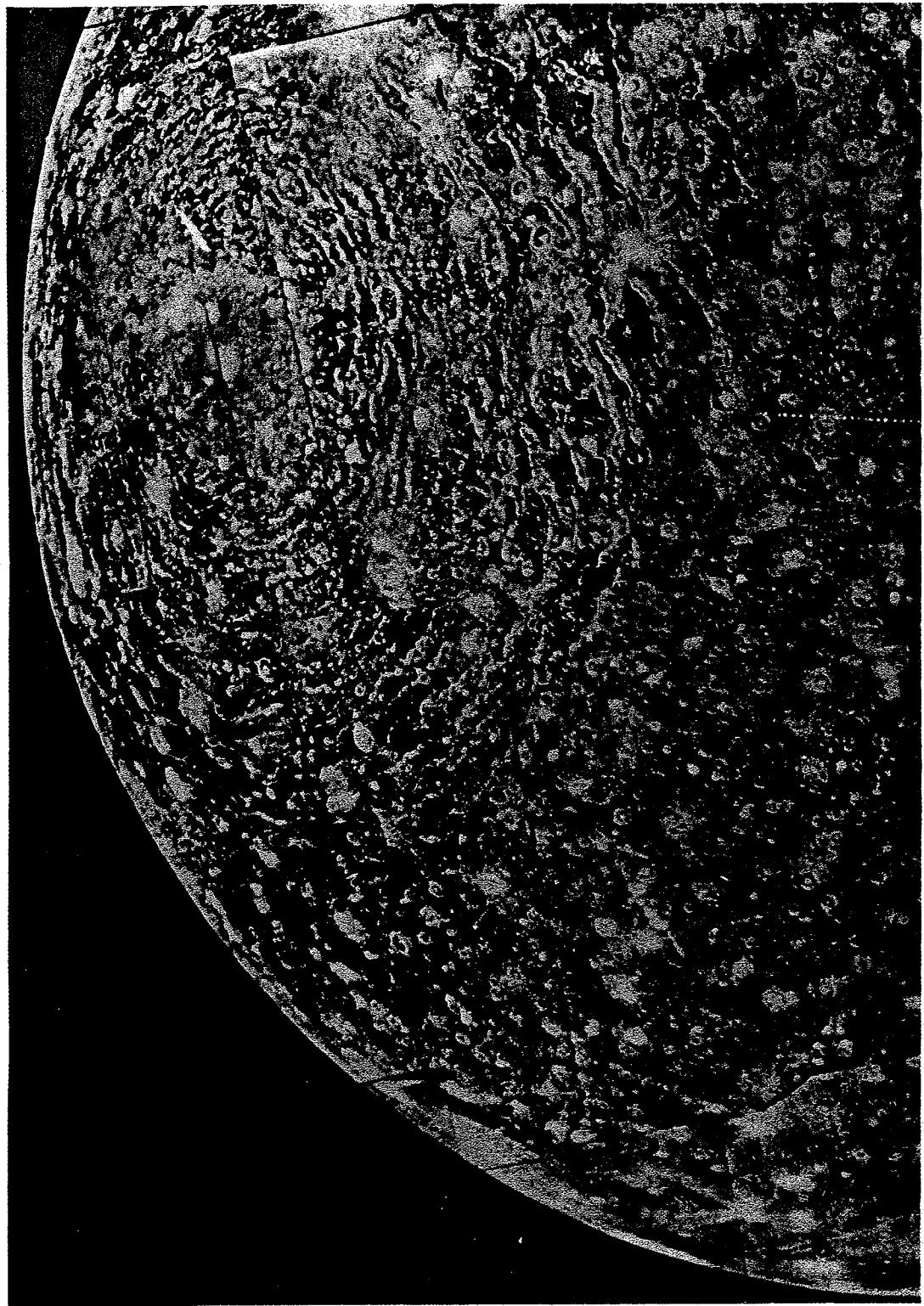
Having thus warned the reader of differences between this book and traditional approaches, I shall describe the morphology of the Orientale and Valhalla basins. More detail on the structure of these and other basins will be found in Chapter 9.

Extremely large impact structures on the moon and the icy satellites Ganymede and Callisto seem to have affected the entire tectonic framework of each body. The center of the young Orientale Basin (Fig. 2.7) is located on the moon's western limb at about 20° South. The relatively flat inner floor lies 2 to 3 km below the distant surrounding plains. The basin is encircled by a well-defined inward-facing scarp 2 to 7 km high and about 900 km in diameter. This scarp sharply defines the inner edge of the Cordillera Mountain ring. Within the Cordillera escarpment, a second mountainous ring, named the Rook Mountains, has a diameter of about 600 km and is also bounded on the inward side by a 2- to 7-km high scarp. The Rook escarpment is not as well defined as the Cordillera, and both rings seem to break down and become diffuse in the poorly imaged western quadrant. The terrain between the Rook and Cordillera escarpments slopes gently outward from the crest of the Rooks to the foot of the Cordillera. Two other rings of irregular mountain peaks occur within the Rooks, with diameters of about 480 and 320 km. The ratios of adjacent ring diameters for Orientale and other similar basins are said to be close to  $\sqrt{2}$ , although there is considerable scatter in this relation.

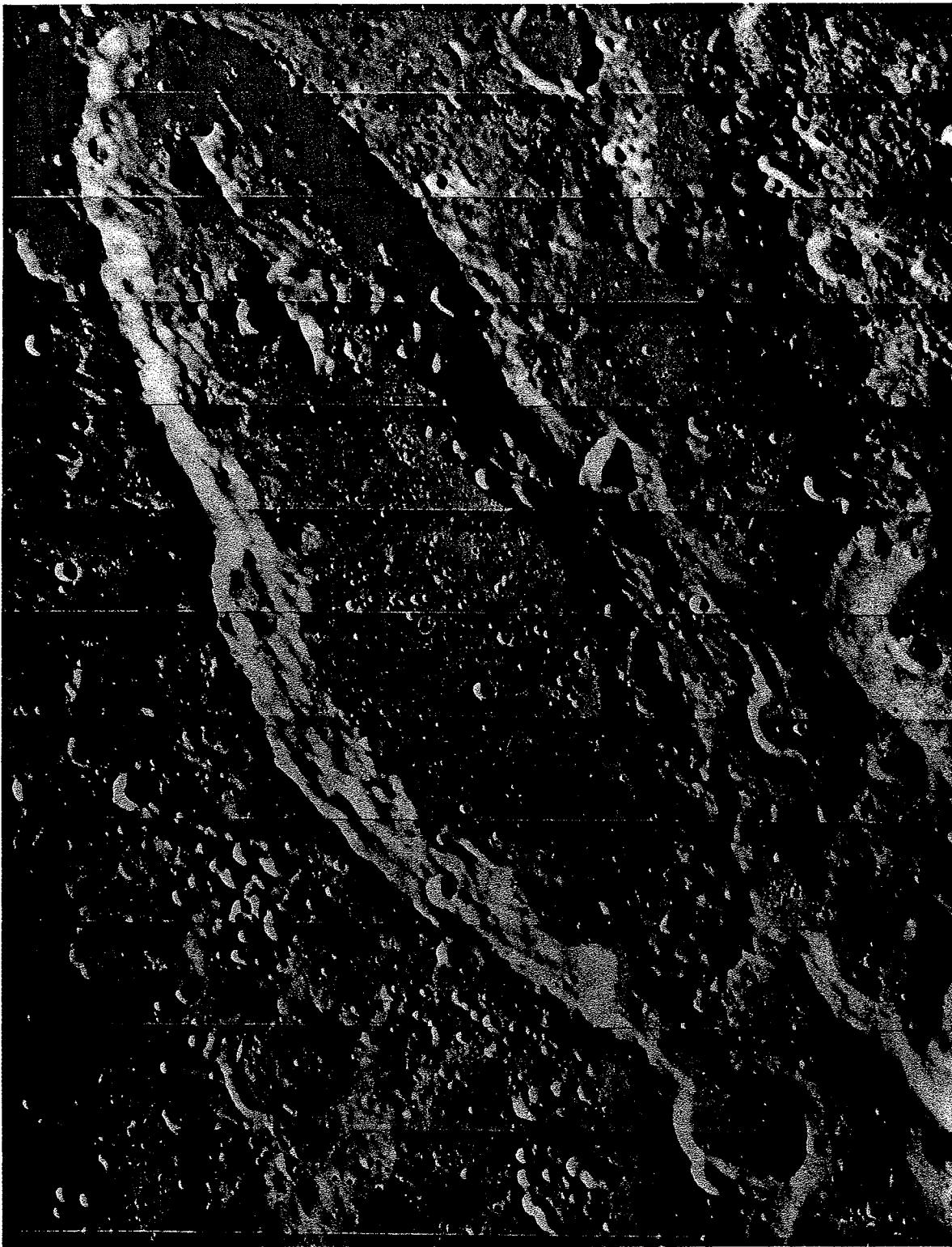
The inner basin of Orientale is 4 to 8 km below the crests of the Rook and Cordillera escarpments. The inner basin is partially flooded with mare basalt, but otherwise has a fissured, hummocky floor, which is interpreted as solidified impact melt mixed with rock debris. Although the Orientale basin is generally believed to be of impact origin, no one has yet succeeded in demonstrating unequivocally which, if any, of the rings corresponds to the rim of the original impact crater. Most authors feel that the original rim coincides with either the Rook or Inner Rook rings. The dramatic Cordillera escarpment must thus have formed *outside* the original crater. Modification of the crater after its excavation was rapid, since ejecta draping the Cordillera indicate that the escarpment was already in existence when the ejecta arrived some 12 minutes after the impact. Like other fresh craters, Orientale is associated with a negative free-air gravity anomaly (a small positive anomaly near its center is correlated with the mare fill). The anomaly is not as large as might be expected by extrapolation from smaller craters, however, implying at least partial isostatic compensation.



**Fig. 2.7** Oblique view of Mare Orientale on the moon. Note the prominent circular inward-facing scarps that surround a small central mare. The rings become somewhat disorganized near the terminator. The deep troughs in the lower part of the photo are features of the ejecta blanket. *Lunar Orbiter IV 194 M.*



**Fig. 2.8** The Valhalla basin on Callisto. The original impact was within the central bright patch, which may represent ejecta. This central zone is surrounded by an annulus of sinuous ridges, which in turn is surrounded by an annulus of outward-facing scarps that can be recognized up to 2000 km from the basin center. *Mosaic prepared by the Jet Propulsion Lab's Image Processing Laboratory.*



**Fig. 2.9** This elongate 100-km long lunar crater named Schiller has a raised rim, central ridge, terraced walls, and flooded floor. It may have been produced by the low-angle impact of one or more projectiles. Alternatively, it may be a volcano-tectonic depression similar to the Toba depression in Sumatra, although this seems unlikely because most of Schiller's morphologic features can also be seen in a crater near Ceraunius Tholus on Mars. That Martian crater also possesses a well-defined ejecta blanket and is clearly of impact origin in spite of its proximity to a shield volcano. Horizontal stripes are 17 km wide. *Lunar Orbiter IV 154 H3.*

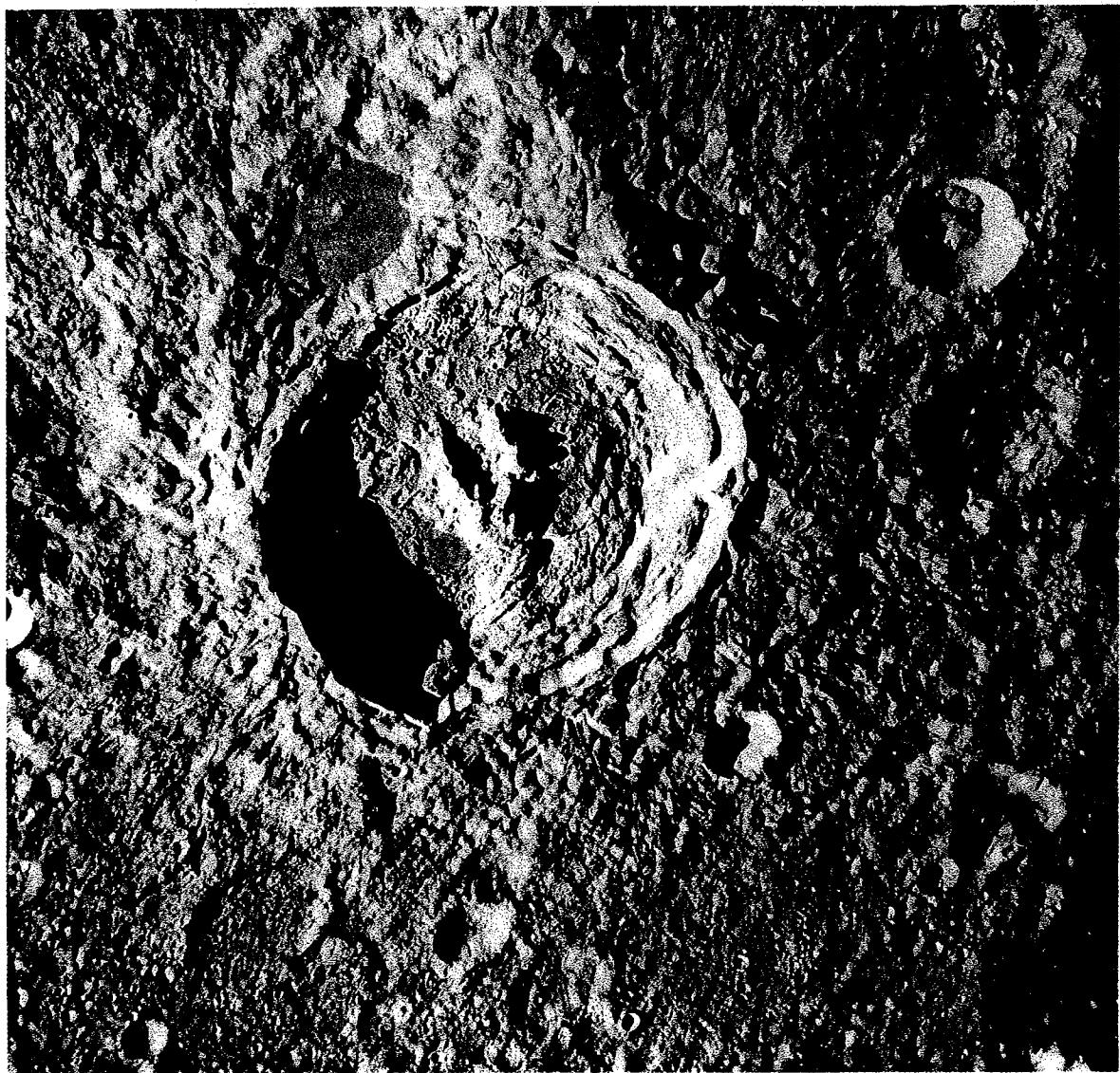
## IMPACT CRATERING

The overlapping ejecta blankets of large craters are the basis of lunar stratigraphy. Orientale's ejecta blanket is recognized as the Hevelius formation in lunar nomenclature, and is divided into several facies. An enigmatic knobby deposit lies between the Rook and Cordillera escarpments. This deposit is generally bounded by the Cordillera escarpment, although lobes of the material locally transgress the scarp. Beyond the Cordillera, the Inner Facies of the Hevelius formation is characterized by radial elongate ridges and troughs. Occasional flow lobes, leveed channels, and transverse ridges within craters (deceleration dunes?) suggest that the deposit originated as some kind of ground-hugging flow. This facies ex-

tends 300 to 600 km from the Cordillera scarp, where it grades into the Outer Facies, a unit that is characterized by weakly lined terrain. Secondary crater chains and clusters occur beyond about 1000 km from the Cordillera scarp.

Although the Orientale-type multiring basin is characterized by an inward facing scarp that appears to have formed outside the original crater, the rings about Valhalla-type multiring basins are either graben or outward-facing scarps that formed at great distances from the original crater. Nevertheless, it is argued in Chapter 9 that the two types of basin have much in common.

The Valhalla Basin on Callisto (Fig. 2.8) is a large and highly complex structure that occupies



**Fig. 2.10** The 75-km diameter crater King on the lunar farside exhibits several aberrant features, the most obvious of which is the distinctive "lobster claw" central peak complex. A massive landslide has formed on the southeast rim (right-hand side) in an area that was topographically high before the impact occurred. A smooth-surfaced pool of impact melt near the top left probably fills a preexisting crater that is mantled by King ejecta. *Apollo Hasselblad photo AS16-122-19580 (H).*

a major fraction of one of Callisto's hemispheres. The outer rings of this basin are located some 2000 km from its center, a distance nearly equal to the 2420-km radius of the satellite. The basin's interior is occupied by a roughly circular patch of bright material 600 to 800 km in diameter. This is surrounded by darker terrain that is traversed by bright, roughly concentric ridges. These ridges have an irregular "ropy" or sinuous appearance similar to lunar mare ridges. They occupy an annulus surrounding the basin center, extending 200 to 300 km outward from the edge of the bright interior. Beyond this zone of concentric ridges is a region of dominantly outward-facing concentric scarps. These scarps are best developed in an annulus about 600 km wide, but individual scarps (especially to the northeast of the basin center) can be seen as far as 2000 km from the center. Near the inner edge of this zone a few inward-facing scarps and graben occur. A light-colored hummocky plains unit appears to have been extruded at the foot of many of these escarpments. The crater density is relatively low on the dark terrain within about 300 km of the edge of the bright basin center, suggesting obliteration of the original crater population by ejecta from the basin-forming impact. Farther away from the basin center the escarpments cut older, heavily cratered terrain. Numerous examples of craters cut by the escarpments may be observed. The down-dropped portions of several such craters appear to have been inundated by the light hummocky plains unit (see Fig. 9.4).

## 2.6 ABERRANT CRATER TYPES

Close scrutiny of any planetary landscape will usually reveal some craters that do not fit the size-morphology progression described here. Most of the bizarre craters that have thus turned up can be explained as a result of special conditions of the impact or the target in which they formed. For instance, most elliptical craters are due to highly oblique impact (Fig. 2.9), layers of different strength in the target produce concentric craters, strong regional joint trends may result in square or polygonal craters, and preexisting topography may produce extra-wide terraces in the walls of complex craters adjacent to topographic highs (Fig. 2.10). Craters on the moon and Mercury are broadly similar, but many Martian craters have peculiar fluidized ejecta blankets with petal-like lobes and terminal ramparts. Other Martian craters seem to sit on pedestals, or have enormous conical central peaks. Some of these phenomena are believed to be the result of a volatile, probably liquid water, not far below the surface. Craters on Ganymede and Callisto frequently possess large central pits, some of them with rims. The origin of these pits is still argued. Geologic studies of terrestrial craters a few kilo-

meters in diameter suggest that their original form was that of an inverted sombrero, with a broad shallow moat surrounding a deeply disturbed central zone. Craters of this type may be due to fragmentation and dispersion of the incoming projectile by aerodynamic stresses, so that by the time it strikes the ground the projectile is more similar to a tightly clustered shotgun blast than a solid chunk of rock.

These and other aberrant crater types will be discussed and illustrated in the appropriate parts of succeeding chapters. It is enough here to mention them and to indicate that, like all simple descriptive schemes, the size-morphology progression has exceptions and caveats. Nevertheless, the size-morphology progression is surprisingly useful in organizing the fundamental forms of impact craters over a wide range of diameter scales.

## FURTHER READING

### General References

The most complete description of lunar craters and other surface features is in a remarkable book of annotated Lunar Orbiter photographs, all printed at very high quality:

P. H. Schultz, *Moon Morphology*, University of Texas Press, Austin and London, 626 pp. (1976).

A more recent collection of annotated photographs of craters and other lunar features from the Apollo missions is

H. Masursky, G. W. Colton, and F. El-Baz, *Apollo Over the Moon: A View from Orbit*, NASA SP-362, U.S. Government Printing Office, Washington, D.C., 255 pp. (1978).

No work comparable to Schultz's has appeared for any other planet, although many Martian craterforms are well illustrated in

M. H. Carr, *The Surface of Mars*, Yale University Press, New Haven and London, 232 pp. (1981).

### Special Topics

The special morphology of microcraters on lunar rock specimens is clearly treated by the following paper and references therein:

F. Hötz, J. B. Hartung, and D. E. Gault, Micrometeorite craters on lunar rock surfaces, *J. Geophys. Res.* 76: 5770-5798 (1971).

The size-morphology sequence for lunar craters first described in Gilbert (1893) has been refined and extended by many subsequent authors. The modern view of this sequence was established in an interesting exchange of papers. The first paper was that of

E. I. Smith and A. G. Sanchez, Fresh lunar craters: Morphology as a function of a diameter, a possible criterion for crater origin, *Modern Geology* 4: 51-59 (1973).

## IMPACT CRATERING

R. J. Pike objected to Smith and Sanchez's emphasis on the continuity of the sequence in

R. J. Pike, Size-morphology relations of lunar craters: a discussion, *Modern Geology* 5: 169-173 (1975).

Smith and Sanchez concurred with Pike on the existence of a simple-to-complex discontinuity but not on the diameter at which it occurs:

E. I. Smith and A. G. Sanchez, Fresh lunar craters: Morphology as a function of diameter, a possible criterion for crater origin: Reply, *Modern Geology* 5: 175-176 (1975).

The entire question of lunar crater morphology is well reviewed by

K. A. Howard, Fresh lunar impact craters: Review of variations with size, *Proc. 5th Lunar Sci. Conf.* (Supplement *Geochim. Cosmochim. Acta*) 1, 61-69 (1974).

Craters on Mars and Mercury are similar to those on the moon, except that the transition diameters differ, as described by

M. J. Cintala, J. W. Head, and T. A. Mutch, Characteristics of fresh Martian craters as a function of diameter: Comparison with the moon and Mercury, *Geophys. Res. Lett.* 3: 117-120 (1976).

Note that several controversies have developed over the transition diameters, especially for Mercury. The latest work on this issue is described in Chapter 8.

The best general geologic description of Meteor Crater is still

E. M. Shoemaker, Impact mechanics at Meteor Crater, Arizona, in B. M. Middlehurst and G. P. Kuiper (Eds.), *The Moon, Meteorites and Comets* 4: 301-336, University of Chicago Press, Chicago and London, 810 pp. (1963).

The structure and ejecta deposits of lunar multiring basins, Orientale in particular, are well described by

K. A. Howard, D. E. Wilhelms, and D. H. Scott, Lunar basin formation and highland stratigraphy, *Rev. Geophys. Space Phys.* 12: 309-327 (1974).

and

H. J. Moore, C. A. Hodges, and D. H. Scott, Multi-ringed basins—Illustrated by Orientale and associated features, *Proc. 5th Lunar Sci. Conf.* (Supplement *Geochim. Cosmochim. Acta*) 1: 71-100 (1974).

### III. STRESS WAVES IN SOLIDS

#### 3.1 INTRODUCTION

An understanding of the process of impact cratering requires a basic knowledge of how stress waves originate, propagate, and decay. The initial impulse delivered by the impactor is carried to the target by stress waves of varying strength and direction of motion. The crater's excavation—its ultimate diameter, depth, and shape—are critically dependent upon the interaction of stress waves with the free surface as well as upon the thermodynamics of the stress wave itself. The petrological and mineralogical changes induced by impacts, such as impact melting or phase transformations, depend upon shock pressure and duration.

This chapter treats those aspects of stress waves in solids that are most important in impact cratering. The description is one-dimensional: spherically expanding waves are discussed in Chapter 4. Pressure (sound) waves and weak elastic waves in solids are reviewed first, then the more specialized topic of strong stress waves are discussed. Topics not usually discussed in introductory texts, such as thermodynamic irreversibility in stress waves, plastic yielding, and the effects of porosity on wave propagation, are included here because they are important for crater excavation. Historically, numerical computation of impact or explosion crater sizes was considered difficult and unreliable. It now appears that the difficulty was caused by the neglect of these "second-order" phenomena and the use of overly simplified equations of state. The velocity of the excavation flow field is very sensitive to irreversible processes in stress waves. These topics will be explored here and further in Chapter 5.

#### 3.2 PRESSURE WAVES IN FLUIDS

The most familiar type of stress wave is the pressure, or sound, wave. It is a progressive wave that travels through homogeneous fluids (liquids or gases) at a constant speed. The amplitude of the wave can be described as a pressure fluctuation above or below the mean, a density fluctuation, or as a particle velocity variation. Figure 3.1 illustrates these factors for a rectangular pulse.

*Particle velocity* is the velocity of a small area of the medium (much smaller than a wavelength but much larger than an atom) that is alternately

accelerated and decelerated as the wave passes. Particle velocity is much slower than the propagation velocity in weak waves, and even in the strongest waves must be less than the wave speed. The direction of particle velocity is longitudinal in a pressure wave; that is, the velocity vector is parallel to the direction of propagation.

The equation describing a one-dimensional pressure wave can be written either in terms of pressure,  $P$ , or longitudinal wave velocity,  $u_L$ . It involves two partial derivatives in both space and time. Equations of this form are known as *wave equations* since their solutions include disturbances (either wave trains or pulses) that travel at constant speed through the medium. These equations are:

$$\frac{\partial^2 P}{\partial t^2} = c_B^2 \frac{\partial^2 P}{\partial x^2} \quad (3.2.1)$$

$$\frac{\partial^2 u_L}{\partial t^2} = c_B^2 \frac{\partial^2 u_L}{\partial x^2} \quad (3.2.2)$$

where  $x$  is distance along the direction of propagation,  $t$  is time, and  $c_B$  is the wave speed. This speed is computed from the zero-pressure isentropic bulk modulus (inverse compressibility) of the medium  $K_0$  and its uncompressed density  $\rho_0$  by

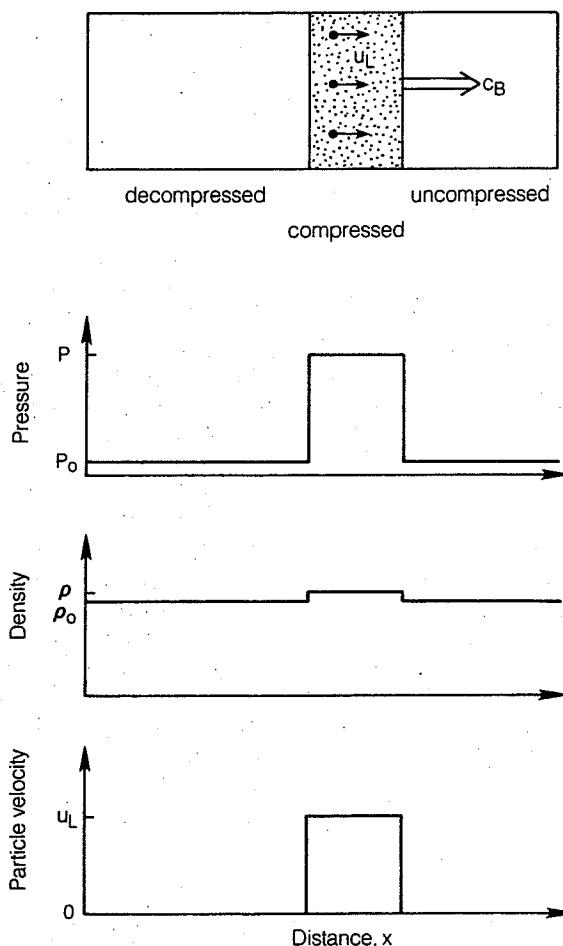
$$c_B = (K_0/\rho_0)^{1/2} \quad (3.2.3)$$

The pressure and particle velocity are directly proportional to one another in a weak pressure wave:

$$P = \rho_0 u_L c_B \quad (3.2.4)$$

Equations 3.2.1 and 3.2.2 are linear, so that weak pressure waves can be superposed: the pressures or particle velocities of two or more waves passing the same spot can be linearly or vectorially added, respectively. The stress state in the medium is fully described by the pressure,  $P(x, t)$ , at every point  $x$  in the medium at some time,  $t$ . There are no shear or differential stresses in a pressure wave.

The energy density in the wave is the sum of the kinetic energy per unit mass ( $1/2$ )  $u_L^2$ , and the work of compression per unit mass,  $PdV$ , integrated from the wave's arrival time.  $V$  is the volume per unit mass ( $dV = -d\rho/\rho_0^2$ ). These two terms are equal in a freely propagating wave or pulse.



**Fig. 3.1** The major features of an elementary pressure wave. The wave travels at constant speed  $c_B$  through a fluid medium. Pressure  $P$ , density  $\rho$ , and particle velocity  $u_L$  all increase in the compressed region of the wave, then return to their initial values after the wave passes.

The next section generalizes the concept of pressure waves to weak elastic waves in solids.

### 3.3 WEAK WAVES IN SOLIDS

#### 3.3.1 Elastic waves

Elastic waves are more complex than pressure waves because, unlike gases or liquids, elastic materials can support differential stresses. Stress waves in homogeneous solids are of two principal types: longitudinal and transverse. Longitudinal waves are similar to the pressure waves just discussed; transverse waves have no analog in fluids. Transverse waves arise because solids, besides resisting compression, also resist change in shape or distortion. This resistance is expressed by the shear modulus,  $\mu$ , which is almost always smaller than the bulk modulus  $K_0$ . Transverse waves thus travel more slowly than longitudinal waves.

These waves are called "transverse" because the direction of motion of a small area of the solid

accelerated by the wave is perpendicular to the direction of wave propagation. There are actually two independent transverse waves whose particle velocity vectors are orthogonal to one another while both are perpendicular to the propagation direction. Both transverse waves travel at the same speed in homogeneous solids, but they are reflected with different amplitudes from interfaces between different substances or from free surfaces. They may travel at different speeds in non-isotropic media.

Elastic waves are described by linear wave equations similar to Equation 3.2.2. The pressure alone, however, does not adequately describe the stress state in the wave. The equations are therefore written in terms of the particle velocity vector, whose three components can describe the full panoply of stress waves in solids. In one dimension (for a plane wave of infinite extent perpendicular to the direction of propagation) the equations of motion are:

$$\frac{\partial^2 u_L}{\partial t^2} = c_L^2 \frac{\partial^2 u_L}{\partial x^2} \quad (3.3.1)$$

$$\frac{\partial^2 u_{T1,T2}}{\partial t^2} = c_T^2 \frac{\partial^2 u_{T1,T2}}{\partial x^2} \quad (3.3.2)$$

where  $u_L$  is the particle velocity in a longitudinal wave that propagates at speed  $c_L$ , and  $u_{T1}$  and  $u_{T2}$  are orthogonal transverse particle velocity components that both propagate at speed  $c_T$ . The propagation speeds depend upon the bulk and shear moduli:

$$c_T = (\mu/\rho_0)^{1/2} \quad (3.3.3)$$

$$c_L = [(K_0 + 4\mu/3)/\rho_0]^{1/2} \quad (3.3.4)$$

The longitudinal wave speed in a solid (Eq. 3.3.4) is greater than the speed in a fluid (Eq. 3.2.3) because the solid's resistance to distortion, parameterized by  $\mu$ , augments its resistance to compression, parameterized by  $K_0$ . The transverse wave speed (Eq. 3.3.3) depends only upon the shear modulus  $\mu$  because transverse wave motion does not cause a change in the volume of the material it passes through.

The stress state in an elastic wave must be described by a number of different stress components in addition to the pressure. Longitudinal stresses, transverse stresses, and various shear stresses collectively form the components of a tensor. Even in a plane longitudinal wave, the stress must be represented by a tensor, since it involves both distortion and compression, making the stress components parallel and perpendicular to the direction of propagation unequal. These stresses are given by equations analogous to (Equation 3.2.4):

$$\sigma_L = -\rho_0 u_L c_L \quad (3.3.5)$$

$$\sigma_P = \left( \frac{\nu}{1-\nu} \right) \sigma_L \quad (3.3.6)$$

where  $\sigma_L$  is the longitudinal stress,  $\sigma_T$  is the stress component perpendicular to the propagation direction (there are actually two orthogonal but equal components), and  $\nu$  is Poisson's ratio, a dimensionless material constant given by

$$\nu = \frac{1}{2} (3K_0 - 2\mu) / (3K_0 + \mu) \quad (3.3.7)$$

Note the minus sign in Equation 3.3.5: in this book stresses are presented according to the "engineering convention," in which positive stresses are tensional, negative stresses compressional.

The stress in a transverse wave is pure shear (the only nonzero components of the stress tensor are off diagonal). The magnitude of this shear is

$$\sigma_S = \rho_0 u_T c_T \quad (3.3.8)$$

The magnitude of the stress in a transverse wave is generally less than that in a longitudinal wave, given similar particle velocities, because the transverse wave speed is generally smaller than the longitudinal wave speed.

The energy density in either of these wave types is, like that in a pressure wave, equal to the sum of the mean kinetic and distortional energies over a wave cycle. A fraction of the energy imparted by an impact is ultimately carried away by longitudinal and transverse elastic waves. In some cases this energy may be refocused in areas distant from the impact site, causing rapid geological modifications. (See Sec. 5.2.6 for a discussion of the seismic energy produced by an impact and Section 12.6 for a discussion of impact-induced seismic modification of landforms.)

Transverse waves (and surface waves, not discussed here) are probably not of much importance for the cratering process. This is because the strength of such waves is limited by the shear strength of solids, whereas the strength of longitudinal waves has no limit. The stresses developed during the early stages of a hypervelocity impact are so much larger than the strength of known solids that transverse waves may be neglected. This approximation fails in the later phases of crater excavation, but may not have much effect on the final crater.

### 3.3.2 Reflection of elastic waves at interfaces and free surfaces

Impact craters form near the free surface of targets that are often composed of layers with contrasting elastic properties. The interaction of impact-generated stress waves with free surfaces and interfaces is thus an important part of cratering. This interaction can be treated exactly only for linear waves, but the concepts thus derived are generally useful. This subject is discussed with great clarity by Rinehart (1960). The present section describes only the major conceptual results necessary for an understanding of impact crater-

ing and the discussion is therefore confined to longitudinal waves.

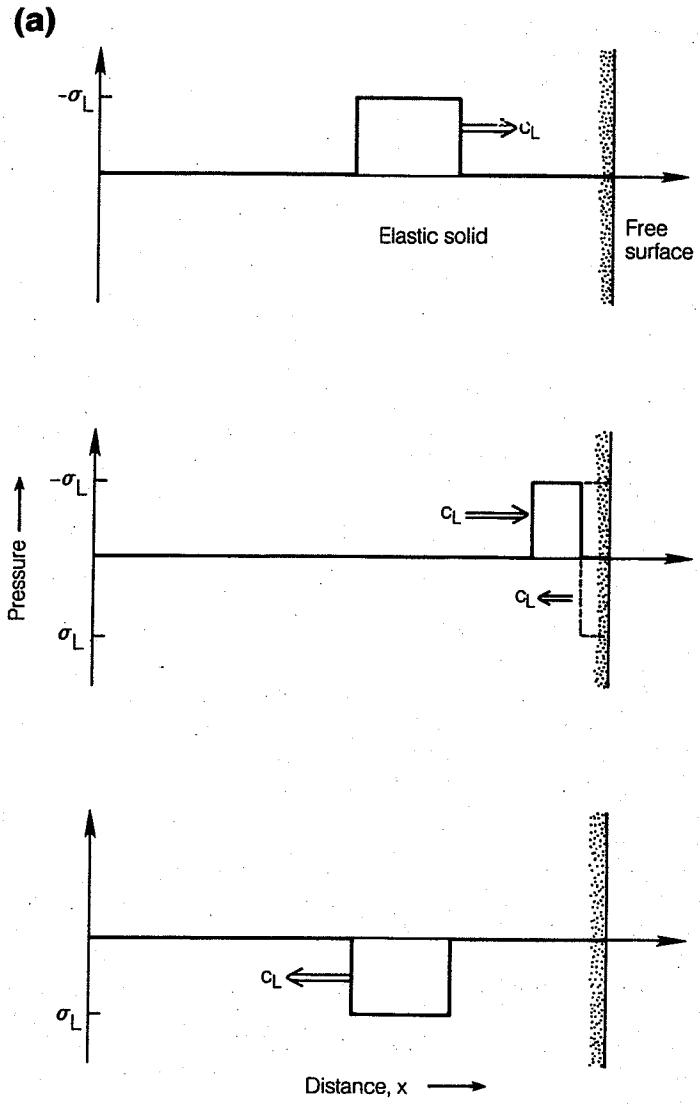
Reflection at interfaces is particularly simple in elastic waves because they are linear: the stress or particle velocity in an area affected by two such waves is a component-by-component sum of the stresses or particle velocities of each wave. Since the stronger plastic, or shock, waves (described in the next sections) cannot be linearly superposed, they must be analyzed numerically. Nevertheless, the qualitative understanding gained in the study of elastic waves is generally applicable to these nonlinear waves as well.

Any wave, elastic or not, that impinges on an interface where material properties change suddenly must preserve the continuity of certain physical quantities. Thus, the mutual impenetrability of solids at a contact requires the same particle velocity in the adjacent substances. (An exception occurs if a crack opens at an interface under extension; such cases require special consideration.) Similarly, the absence of external forces at a contact requires continuous stresses, both normal and shear, across the contact surface. These boundary conditions are satisfied by the creation or modification of elastic waves at the contact.

Figure 3.2a illustrates a rectangular compressional pulse approaching a free surface, and its reflection as a tensional pulse. By definition, a free surface cannot support normal or shear stresses. When a pulse strikes it, a new elastic wave is created, maintaining the normal stress at zero (there are no shear stresses in this example). Because elastic waves in one dimension must travel either to the left or to the right with speed  $c_L$ , and because a rightward-moving pulse generated at the surface could not influence the stress in the elastic medium, a leftward-moving pulse therefore arises at the surface. The zero normal-stress boundary condition requires a tensional pulse, so that the sum of the longitudinal stresses  $\sigma_L$  in the two pulses is zero. Once generated at the boundary, the tensional pulse propagates to the left without modification because there is no other source of elastic waves in the medium. The rightward-moving compressional pulse and the leftward-moving tensional pulse do overlap for a short period of time, giving rise to a locally complex stress history, but since the waves are linear the net stress is merely the sum of the stresses in the two pulses at any given time.

Although the normal stress must vanish on a free surface, the particle velocity does not. The vacuum to the right of the elastic medium offers no resistance to its motion, which is thus determined by the sum of the particle velocities in the compressional and tensional pulses. The particle velocity in the compressional pulse, given by Equation 3.3.5, is  $u_L$  to the right. The particle velocity in the tensional pulse, however, is also  $u_L$  to the right—a result of the reversed signs of both

## IMPACT CRATERING



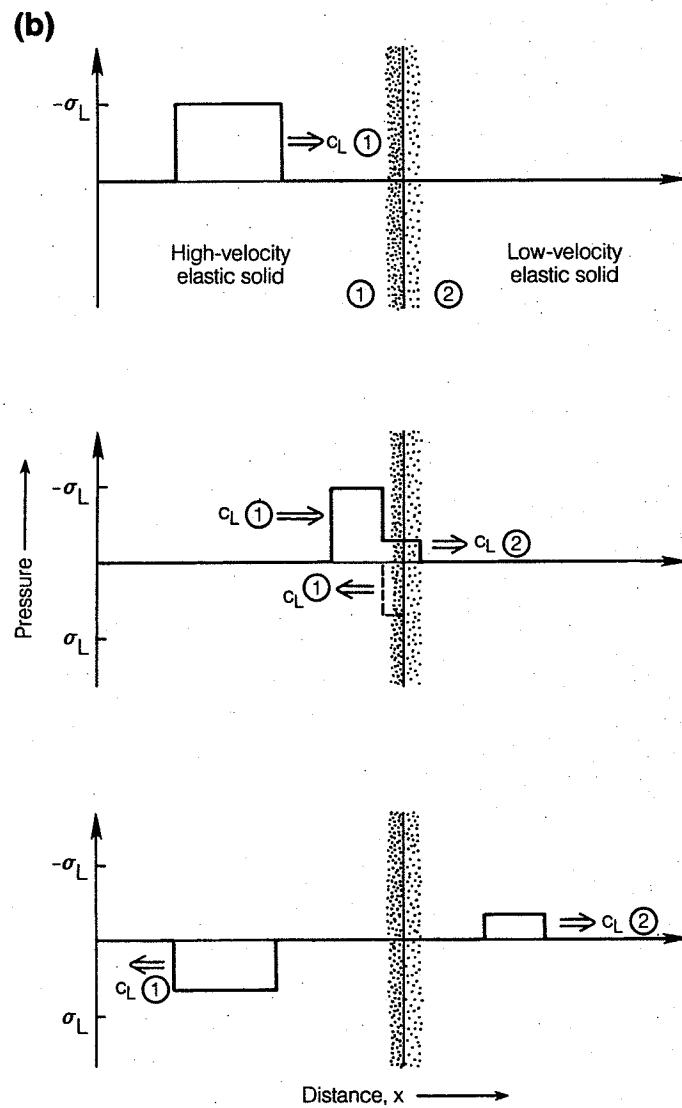
**Fig. 3.2a** Reflection of an elastic wave from a free surface. Since the net pressure at a free surface is zero by definition, the initially compressive wave is reflected back into the solid as a tensile wave. Conversely, an initially tensile wave would be reflected as a compressive wave. The particle velocities in both waves point toward the right. For a short time during reflection the free surface moves rightward with twice the particle velocity of the incident wave. If the wave is strong enough the solid fails under the tensile stress and a "spall" flies off to the right at a speed of  $2 u_p$ .

the stress and direction of motion. The net particle velocity during the time that the pulses are interfering is thus  $2u_p$  to the right.

This result is known as the "velocity doubling rule." The particle velocity at free surfaces is approximately twice that in the incident pulse even for strong stress waves. At one time this rule formed the basis of a technique for measuring particle velocities: fragments called *spalls* often break off near the free surface. The velocity of these spalls is roughly twice the particle velocity in the original compressional pulse, which is thus inferred from the spall velocity by dividing it by two. The spall process and its importance for secondary craters are discussed in Sections 5.4 and 6.4. Although it is only of slight importance in

crater excavation because of the small quantity of material involved, spall is responsible for ejecting some solid target material at speeds that may exceed planetary escape velocities.

Figures 3.2b and 3.2c illustrate the interaction of a compressional rectangular pulse moving to the right with (b), an interface where the wave velocity is smaller on the right than on the left, and (c), an interface where the wave velocity is larger on the right. Case (b) is similar to wave reflection at a free surface, except that the continuity of particle velocity as well as normal stress across the contact requires the creation of a third pulse, a rightward-moving compressional pulse that travels into the elastic material to the right of the interface. A tensional pulse is still reflected into the



**Fig. 3.2b** Reflection of a pressure pulse that encounters an interface between high- and low-velocity material. Part of the wave is reflected back into the high-velocity material as a tensile wave, as in Figure 3.2a, while the rest of the wave continues into the low-velocity material.

medium at the left. Because the energy of the original pulse is now divided between the transmitted compressional pulse and the reflected tensional pulse, the reflected pulse is less intense than it was when the contact was a free surface.

In case (c) the pulse travels across an interface from a low-sound-velocity medium into a high-sound-velocity medium. As before, there is a transmitted wave and a reflected wave, but in this case *all* the pulses are compressional. In the limit of infinite sound velocity on the right, the interface approximates a rigid wall that reflects the incident wave perfectly. In this limit the particle velocity at the wall is zero and the maximum stress is twice that in the incident pulse.

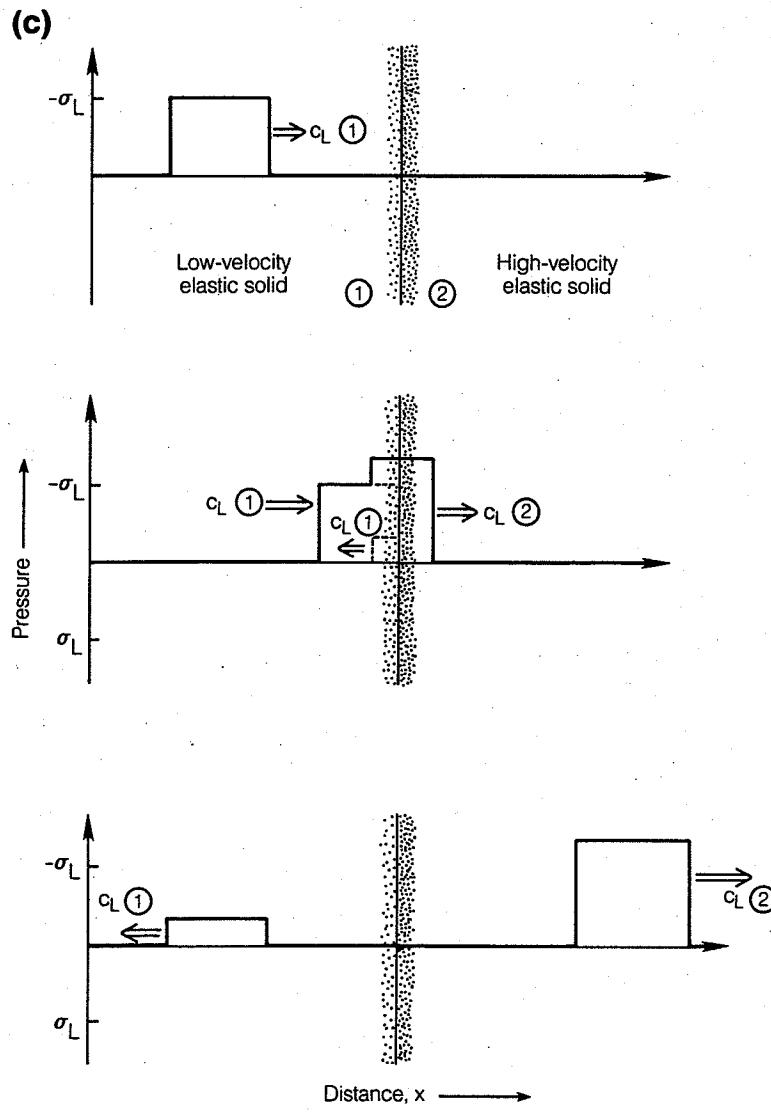
The examples described so far all treat the behavior of a rectangular compressional pulse. The same results apply to any pulse shape or sign (of course, if the sign of the incident pulse is changed,

the signs of all the resultant pulses must likewise be reversed). Thus, a triangular compressional pulse moving to the right toward a free surface is reflected as a triangular tensional pulse (reversed left-to-right from the original pulse) traveling to the left. The pulse shape is preserved because the stress in the original pulse must be canceled at the free surface at all times. Only a tensional pulse of precisely the same shape as the original is capable of this. Analogous results hold for the other cases where the interface is not a free surface.

### 3.3.3 Plastic yielding at the Hugoniot Elastic Limit

The particle velocity and stress in an elastic wave increase as the strength of an initial disturbance increases. The stress ultimately reaches a limiting value beyond which plastic, or irreversible, dis-

## IMPACT CRATERING



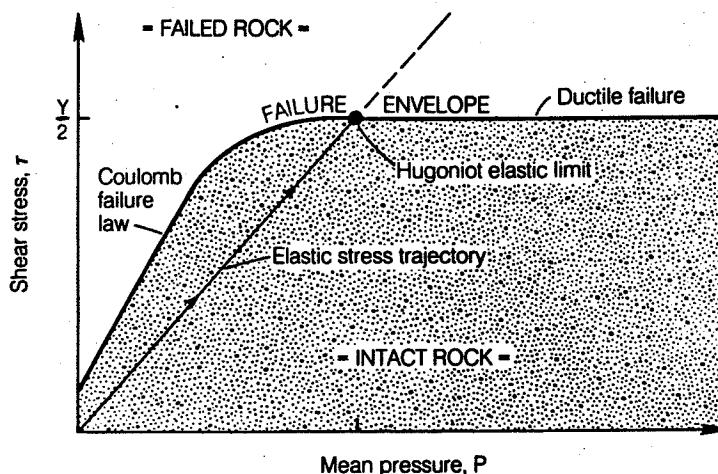
**Fig. 3.2c** Reflection of a pressure pulse that encounters an interface between low- and high-velocity material. The reflected wave is compressional, like the incident wave itself. A compressional wave also continues into the high-velocity material.

tortions occur in the solid medium through which the wave propagates. This plastic yielding affects both the speed and shape of the stress wave. The onset of this behavior is indicated by a characteristic kink in the Hugoniot P-V curve (see Sec. 3.4) at a pressure that is known as the Hugoniot Elastic Limit.

Although solids can resist almost arbitrarily large compressive stresses, their resistance to stress differences is limited. The subject of the strength of solids is a large and complex one. However, for the purposes of this discussion, we will consider only the response of a mass of rock to a longitudinal wave. The longitudinal stress component in an elastic longitudinal wave is a factor  $(1 - \nu)/\nu$  larger than the transverse stress components (Eq. 3.3.6). This factor ranges between 2 and 3 for most rocks. As the strength of the wave increases, the absolute difference be-

tween the longitudinal and transverse stress components likewise increases. Eventually, this stress difference reaches a yield stress,  $-Y$ , plastic flow begins, and little subsequent increase in the stress difference occurs.

Figure 3.3 schematically illustrates the failure surface of a "typical" rock on a plot of maximum shear stress (a function of the difference between the longitudinal and transverse stresses),  $\tau = -(\sigma_L - \sigma_P)/2$ , versus mean pressure,  $P = -(\sigma_L + 2\sigma_P)/3$ . At low mean pressures (less than about 2 GPa) the failure envelope of rocks is described by a Coulomb friction law: the strength rises linearly with increasing mean pressure. At higher pressures, the strength is nearly constant: failure is characterized by plastic distortion of constituent minerals (ductile flow) rather than by growth and coalescence of microcracks (brittle fracture). The slope of the failure envelope in the Coulomb fric-



**Fig. 3.3** Failure of rock in a strong compressive wave. The maximum resolved shear stress at failure,  $\tau$ , is a linear function of total pressure  $P$  in the Coulomb failure regime. The failure envelope is independent of pressure in the high-pressure ductile failure regime. The stress in a strong elastic wave increases along the trajectory shown, remaining within the failure envelope until it intersects the ductile portion at the Hugoniot Elastic Limit.

tion portion of the curve (30 to 45° for most rocks) rises more steeply than the line representing the stress in a longitudinal elastic wave, which begins at the origin and has an upward slope of  $3(1 - 2\nu)/2(1 + \nu)$ —about 30 degrees for  $\nu = 0.25$ , which is typical for rocks (see Fig. 3.3). Failure thus does not occur at low stresses. However, the failure envelope eventually flattens out and the elastic stress trajectory intersects it at  $\tau = Y/2$ . The longitudinal stress at this point is

$$\sigma_L = -\sigma_{HEL} = -\left(\frac{1-\nu}{1-2\nu}\right)Y \quad (3.3.9)$$

where  $\sigma_{HEL}$  is defined as the Hugoniot Elastic Limit.

Beyond this limit, the maximum shear stress  $\tau$  remains constant at  $Y/2$  as  $P$  increases; that is, both longitudinal and transverse stresses increase in such a way that  $-(\sigma_L - \sigma_p)/2$  always equals  $Y/2$ . For extremely strong waves this difference between the longitudinal and transverse stresses is small compared to the mean stress, and is frequently neglected: in this case, the wave can be approximately described as a strong pressure wave.

Table 3.1 is a compilation of Hugoniot Elastic Limits of a number of geologic materials. These values are only intended to be illustrative: the Hugoniot Elastic Limit varies considerably even among rocks of the same type. Further complica-

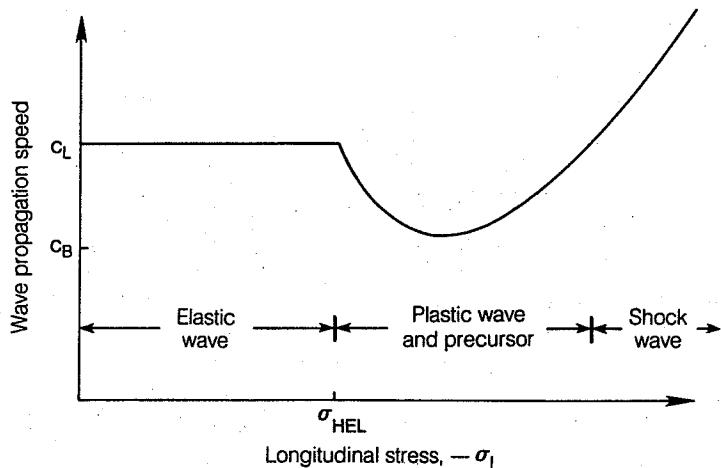
TABLE 3.1 *Hugoniot elastic limits*

Material	Hugoniot Elastic Limit $\sigma_{HEL}$ (GPa)	Source
<i>Single Crystals:</i>		
Periclase ( $MgO$ )	2.5	Grady (1977)
Feldspar	3.	Grady and Murri (1976)
Quartz ( $SiO_2$ )	4.5–14.5*	Duvall and Graham (1977)
Olivine ( $Mg_2SiO_4$ )	9.	Raikes and Ahrens (1979)
Corundum ( $Al_2O_3$ )	12–21*	Grady (1980)
<i>Rocks:</i>		
Halite	0.09	Larson (1982)
Blair Dolomite	0.26†	Larson (1977)
Vermont Marble	0.9	Grady (1977)
Westerly Granite	~ 3	Larson (1977)
Lunar Gabbroic Anorthosite	3.5	Ahrens et al. (1973)
Granodiorite	4.5	Borg (1972)
<i>Metals:</i>		
Armco Iron	0.6	Rice et al. (1958)
SAE 1040 Steel	1.2	Rice et al. (1958)

\*HEL depends upon the crystal orientation.

†Rate dependence observed.

## IMPACT CRATERING



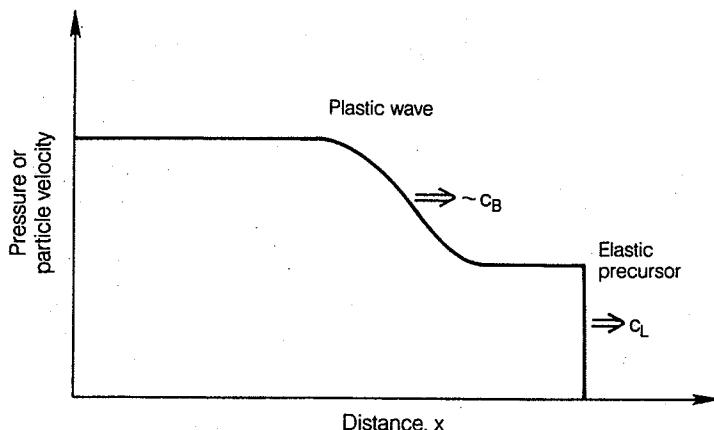
**Fig. 3.4** The propagation speed of a strong compression wave depends upon its strength. Relatively weak elastic waves travel at a constant speed  $c_L$ . Stronger waves break into an elastic precursor and a slower-moving plastic wave that may travel as slowly as the bulk sound speed  $c_B$ . Still stronger waves travel faster than the elastic sound speed and thus have no precursors.

tions result from the presence of porosity, where two more-or-less distinct yield points may be observed, one associated with the collapse of the pores and the other with the onset of ductile flow. Many rocks have rate-dependent failure envelopes without a fixed yield stress. The presence of a fluid phase (usually water) or a soft solid phase (e.g., water ice) within the pores of the rock may also lead to a complicated failure envelope. Individual measurements on actual rock specimens are generally needed to define the failure envelope with adequate precision for numerical computation of crater dimensions, at least for small ("strength-dominated") craters.

The propagation speed of a wave that has exceeded the Hugoniot Elastic Limit drops precipitously. In an elastic wave both the shear and bulk moduli contribute to the longitudinal wave veloc-

ity  $c_L$ . Beyond the Hugoniot Elastic Limit only the bulk modulus contributes significantly, and the speed drops to nearly the bulk wave speed (Eq. 3.2.3). Because of the high mean pressure in a wave that has exceeded the Hugoniot Elastic Limit, and because the bulk modulus increases with pressure,  $K_0$  in Equation 3.2.3 must be replaced by  $K(P)$ . The bulk wave speed in a high-pressure wave is thus greater than that in a low-pressure wave. In extremely strong stress waves, discussed in the next section, this effect is so large that the wave speed rises even above  $c_L$  (see Figure 3.4).

Since the propagation speed of a plastic wave is smaller than that of an elastic wave, all disturbances strong enough to produce a wave that exceeds the Hugoniot Elastic Limit actually produce two waves: an elastic precursor that travels at



**Fig. 3.5** Structure of a strong compression wavefront. The stress in the wave first jumps to the Hugoniot Elastic Limit in the fast-moving elastic precursor, then rises to its final value in the slower-moving plastic wave. In stronger waves the plastic wave travels faster than the elastic precursor, producing a single sharp rise in pressure: a shock front.

speed  $c_L$  and in which the longitudinal stress equals  $\sigma_{HEL}$  and a slower-moving "plastic" stress wave that travels slightly faster than speed  $c_B$ . Figure 3.5 illustrates this "double pulse" structure schematically. Such elastic precursors are frequently observed on particle velocity recordings made near underground nuclear tests. They presumably could be observed near impact cratering events if the appropriate instrumentation were installed.

Plastic waves are inherently nonlinear: they cannot be superposed the way elastic waves can. Consider two plastic waves of equal strength. The transverse and longitudinal stress components differ by a fixed amount,  $Y$ . If the two waves simultaneously pass the same point (perhaps traveling in opposite directions), the transverse and longitudinal stresses would both double, as would their difference if they could be linearly superposed. This would lead to a stress difference of  $2Y$ , well beyond the strength of the material. Thus, plastic waves cannot be linearly superposed. The interaction of such nonlinear waves is complex and must usually be analyzed numerically. Plastic wave reflection and transmission at free surfaces and interfaces is qualitatively similar to that of elastic waves, but quantitative comparisons are not possible. This is especially true when the tensile pulse reflected from a free surface exceeds the tensile strength of the material and fractures form (this may even happen in elastic wave reflection because the tensile strength of a material is usually much smaller than its Hugoniot Elastic Limit).

### 3.4 SHOCK WAVES

#### 3.4.1 Strong compression waves: The Hugoniot equations

The preceding sections have described how progressively stronger initial disturbances produce different kinds of waves. The weakest disturbances produce elastic waves in solids or pressure (sound) waves in fluids. Stronger disturbances plastically deform the solids through which they travel and propagate more slowly than elastic waves. They are invariably accompanied by an elastic precursor. Fluids, lacking shear strength, have no analog of this wave type. Finally, the strongest disturbances produce shock waves—waves that travel faster than elastic waves in the uncompressed material and are therefore supersonic. Figure 3.4 summarizes the different regimes of wave type as a function of strength.

Shock waves, like plastic waves, are inherently nonlinear. Since it travels faster than sound, a propagating shock wave can overrun elastic waves, adding the elastic wave's energy to its own. The double pulse structure of plastic waves is eliminated by the development of a single super-

sonic shock pulse. Shock fronts tend to be abrupt, to the point where they are often represented as a discontinuous jump of pressure, particle velocity, density, and internal energy. This is, of course, a mathematical idealization, but physical shock thickness are often so small as to make a discontinuity an excellent approximation. The thickness of a shock wave in a gas is only about one molecular mean free path—roughly  $10 \mu\text{m}$ . In solids the thickness may be much larger, depending upon viscosity or the operation of other rate-dependent processes. Typical thicknesses range from millimeters in metals to several meters in rocks, where rate-dependent pore closure and crushing thicken the shock front.

The fundamental equations describing abrupt shock fronts were derived by P. H. Hugoniot in 1887. A simple derivation is presented in Appendix I. These equations involve the pressures  $P_0$  and  $P$  in front of and behind the shock front, respectively, the particle velocity  $u_p$  behind the shock, the shock velocity  $U$ , the compressed and uncompressed densities,  $\rho$  and  $\rho_0$ , respectively, and the specific internal energies  $E_0$  and  $E$  (i.e., energy per unit mass) on either side of the shock. These definitions are illustrated in Figure 3.6. The equations are derived from the conservation of mass, momentum, and energy across the discontinuity. Written conventionally, they are

$$\rho(U - u_p) = \rho_0 U \quad (3.4.1)$$

$$P - P_0 = \rho_0 u_p U \quad (3.4.2)$$

$$E - E_0 = (P + P_0)(V_0 - V)/2 \quad (3.4.3)$$

where  $V (= 1/\rho)$  and  $V_0 (= 1/\rho_0)$  are the compressed and uncompressed specific (per unit mass) volumes, respectively. These equations must be applied in the rest frame of the uncompressed material, where the initial particle veloc-

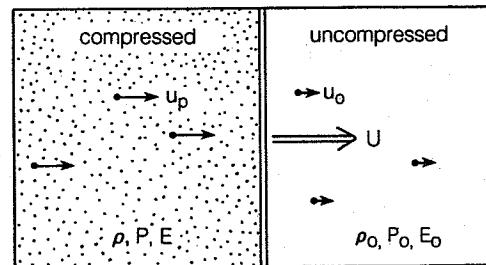


Fig. 3.6 Quantities describing the state of the medium, such as particle velocity  $u_p$ , density  $\rho$ , pressure  $P$ , and internal energy  $E$  (per unit mass), jump discontinuously across a shock front. The Hugoniot equations employ the conservation of mass, momentum, and energy to relate these quantities on either side of the shock front. An equation of state is also necessary to fully define the conditions of the shock. The shock front moves at velocity  $U$ .

ity is zero. If the initial pressure  $P_0$  can be neglected, then  $E - E_0 \approx u_p^2/2$ , which is frequently a useful approximation.

The equations as written are actually valid for both fluids and for the longitudinal stress in a solid, where  $P$  should be replaced by  $-\sigma_L$ . It is *not* correct to replace  $P$  by the scalar pressure in a solid, although at high pressures this distinction can be neglected, since  $P = -\sigma_L - 2Y/3$  and  $P \gg Y$ ; however, at lower pressures, errors can arise unless this point is understood. The above equations are rigorously valid for fluids.

These three equations do not completely specify conditions on either side of the shock: a fourth equation relating the scalar pressure, specific volume, and internal energy is necessary. This equation is generally known as the equation of state:

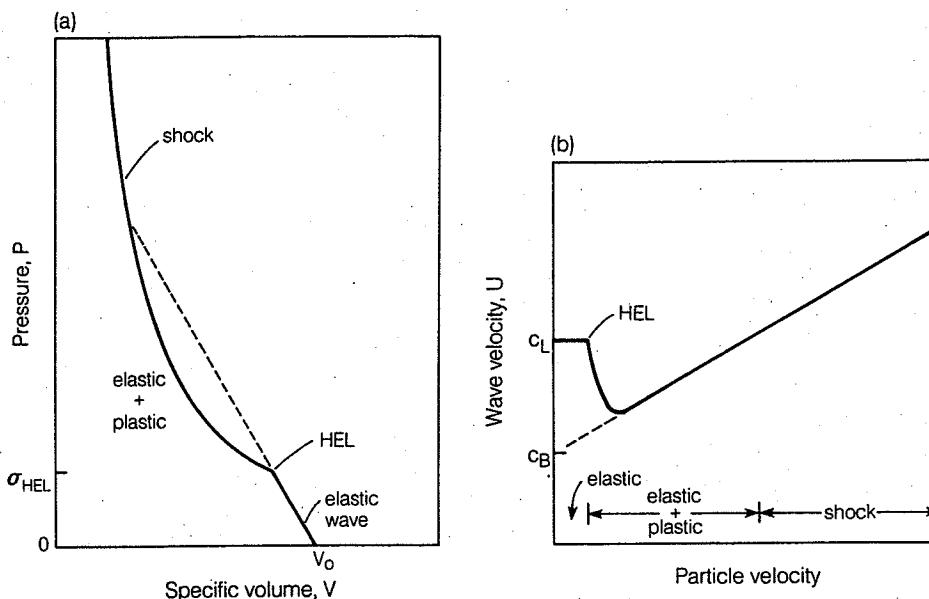
$$P = P(V, E) \quad (3.4.4)$$

The equation of state addresses all of the complexity of the atomic, molecular, and crystalline structure of the material. This equation cannot usually be calculated from first principles, although it can often be approximated—a large amount of effort has been devoted to this task by physicists, and equation of state studies form an important subdiscipline of modern physics. Nevertheless, most of the useful equation of state information in the pressure range of interest for impact cratering has come from experimental investigations.

Equation of state information is gathered by a variety of techniques. The most important tech-

nique for shock wave studies employs an impact between two plates of material, one or both of which may be the material of interest. One of these plates, the "flyer" plate, strikes the "target" plate at high speed, generating a plane shock wave in a manner described in Chapter 4. In one variant of this technique, the flyer plate is accelerated by the detonation of a mass of explosives or by a "gun" in more modern studies. The gun may be a conventional military cannon, a two-stage compression device employing a low-molecular-weight gas such as hydrogen to accelerate a projectile carrying the flyer plate, or, in the most modern developments, an electromagnetic accelerator. No matter what method is used to accelerate the flyer plate, the shock velocity  $U$  and the particle velocity  $u_p$  are measured in the material of interest. These two velocities uniquely determine the equation of state through the Hugoniot equations (3.4.1 to 3.4.3). Unfortunately, current technology cannot produce particle velocities much larger than about 6 km/second, which is smaller than the particle velocities expected in most solar system impacts. Although some data has been obtained at higher velocities using nuclear explosives, the quantities of this data are very limited. The measured equations of state must thus be extrapolated, sometimes by a factor of three or more, before they become useful for the study of natural impact events.

There are several equivalent ways of representing the results of shock wave equation of state studies. Figures 3.7a and 3.7b show two of the



**Fig. 3.7** A material's equation of state may be expressed in terms of either shock pressure  $P$  and specific volume  $V$ , or shock velocity  $U$  and particle velocity  $u_p$ . These alternative descriptions are equivalent through the Hugoniot equations. The figure illustrates the commonly observed linear relation between shock and particle velocities and shows how yielding at the Hugoniot Elastic Limit appears on the plots.

most common:  $P$ - $V$  diagrams and  $U - u_p$  plots. The  $P$ - $V$  diagram is probably the most frequently seen, although the  $U - u_p$  plot has the advantage of representing the quantities actually measured. One coordinate pair ( $P, V$ ) or ( $U, u_p$ ) can be easily converted into the other by use of the first two Hugoniot equations if the uncompressed density  $\rho_0$  is known. For reference, these equations are:

$$V = 1/\rho = (U - u_p)/\rho_0 U \quad (3.4.5)$$

$$P = P_0 + \rho_0 u_p U \quad (3.4.6)$$

and

$$u_p = \sqrt{(P - P_0)(V_0 - V)} \quad (3.4.7)$$

$$U = V_0 \sqrt{(P - P_0)/(V_0 - V)} \quad (3.4.8)$$

where  $V_0 = 1/\rho_0$ .

The third Hugoniot equation (3.4.3) determines the internal energy in the compressed material. Note that it is internal energy, not temperature, that is thus determined. Temperature is difficult to measure in shock wave experiments and must generally be estimated by theoretical methods that are, at best, approximate.

The Hugoniot curves illustrated in Figures 3.7a and 3.7b are frequently misunderstood. The curves do *not* represent a continuum of states that a shocked material passes through as the pressure increases; rather, the curves are the loci of a series of individual shock events. Each point on a curve is the single, unique result of a shock wave compression to the indicated pressure or shock velocity. The shock front is mathematically discontinuous. Quantities change suddenly across it: there are no intermediate states. This point is important because it is easy to confuse the Hugoniot curves on the  $P$ - $V$  plot with a thermodynamic path leading to the high pressure state. Indeed, the release adiabat, which is a thermodynamic path, is often approximated by the Hugoniot curve. However, shock compression is not a thermodynamically reversible process and hence cannot be represented by a continuous path. The Hugoniot equations conserve mass, momentum, and energy, but not entropy.

The  $P$ - $V$  plot has a number of useful features. Equation 3.4.8 shows that the velocity of the shock wave  $U$  for a given shock state is proportional to the square root of the slope of a line connecting the shock state ( $P, V$ ) to the initial state ( $P_0, V_0$ ). The use of this relation is shown in Figure 3.7a where the dashed extension of the elastic portion of the curve separates the high pressure regime of shock waves (strictly, supersonic waves) from that of plastic waves. The dashed line intersects the Hugoniot curve where the shock velocity  $U$  equals the longitudinal wave velocity  $c_L$  (this is most readily seen in Figure 3.7b, but is equally true on both plots since they both represent the same information, although in different coordinate systems), because the slope of the line con-

nnecting the initial and final shock states equals the slope of the elastic portion of the curve, which is also related to the sound velocity by an equation similar to 3.4.8.

Below the Hugoniot Elastic Limit in Figure 3.7a only an elastic wave front is present. A double pulse structure exists along the portion of the Hugoniot curve labeled "elastic + plastic" in Figure 3.7a because the plastic wave travels more slowly than the elastic precursor and thus lags behind it. The velocity of the plastic wave can be found graphically by drawing a straight line from the point labeled "HEL" to the Hugoniot curve at the pressure  $P$  and specific volume  $V$  that characterize the plastic wave. The velocity is computed using an equation similar to 3.4.8 in which  $P_0$  and  $V_0$  are replaced by  $P_{\text{HEL}}$  and  $V_{\text{HEL}}$ . This line necessarily has a smaller slope than the elastic portion of the curve and so represents a wave with lower velocity.

Above the intersection of the dashed line with the Hugoniot curve in Figure 3.7a the speed of the "plastic" wave, given by the slope of the line between the origin ( $P_0, V_0$ ) and ( $P, V$ ), exceeds the elastic wave velocity. The "plastic" wave is thus supersonic and overtakes the elastic wave, producing only one pulse, the shock wave.

From Equation 3.4.3 it follows that when  $P_0$  can be neglected, the specific energy behind the shock,  $E$ , is equal to the uncompressed specific energy,  $E_0$ , plus the area of a right triangle on the  $P$ - $V$  diagram whose hypotenuse is the line connecting the initial and shock states (the Rayleigh line: see Fig. 3.10) and whose sides are parallel to the  $P$  and  $V$  axes.

The principal advantage of the  $U - u_p$  plot shown in Figure 3.7b is that most materials exhibit a linear relation between shock velocity and particle velocity, so long as no phase transformations occur (even in this case, the Hugoniot curve of each phase is often a straight line on the  $U - u_p$  plot). Although there seems to be no fundamental physical reason for it, the equation

$$U = C + S u_p \quad (3.4.9)$$

is often an excellent representation of the shock data up to particle velocities approaching 10 km/second. In this equation, the constant  $C$  is usually close to the bulk sound speed  $c_B$ . The dimensionless slope  $S$  is related to the Grüneisen parameter  $\Gamma$  at low pressure

$$S = (1 + \Gamma)/2 \quad (3.4.10)$$

where  $\Gamma = \alpha K_0 / \rho_0 C_V$ , in which  $\alpha$  is the volume coefficient of expansion,  $K_0$  is the zero pressure isentropic bulk modulus, and  $C_V$  is the specific heat at constant volume. Equation 3.4.9 is the basis of a widely used parameterization of the equation of state called the Tillotson equation (see Appendix II). The Tillotson equation interpolates between Equation 3.4.9 at low pressure (<100 GPa) and

the Thomas-Fermi equation of state at very high pressures ( $> 10,000$  GPa). It is one of the best simple representations of the equation of state in the pressure regime appropriate for natural impact cratering studies.

Figures 3.7a and 3.7b show that there is little fundamental difference between the regime where stress impulses travel as plastic waves (accompanied by an elastic precursor) and the regime where true supersonic shock waves develop. In fact, there is little need to distinguish the two types of stress wave in impact cratering studies. The elastic precursor is too weak to significantly affect the target rocks, and it makes little difference whether the main part of the stress wave moves at subsonic or supersonic velocity. The formal distinction between these two wave types is thus frequently blurred by using the generic term "stress wave" when it does not matter whether the wave is truly supersonic or not.

### 3.4.2 Complexities of the Hugoniot curves

The Hugoniot curves described so far include only one degree of complication beyond that of a structureless fluid medium: the finite strength of a solid produces a cusp in the Hugoniot curve when it is exceeded, and a two-pulse structure develops in which the main stress wave is preceded by an elastic precursor unless the wave is so strong that the plastic wave speed exceeds the elastic wave speed. Other circumstances can also modify the shape of the Hugoniot curve and the propagation of stress waves.

Most crystalline solids undergo structural rearrangements during shock compression, trans-

forming to denser phases at successively higher pressures. Thus quartz, with density  $2650 \text{ kg/m}^3$  at zero pressure, transforms to Stishovite (density  $\rho_0 = 4287 \text{ kg/m}^3$ ) between 15 and 40 GPa. These transformations appear as kinks in the Hugoniot curve. On the  $P$ - $V$  plot a phase transformation to higher density causes an offset of the curve to the left, toward lower specific volume (Fig. 3.8a). Each phase may be considered to have its own Hugoniot curve. The denser phase's curve is sometimes directly observable because the mineral can be recovered as a metastable phase and studied experimentally (Stishovite is an example). The phase transformation appears as a gradual transition from one curve to the other. The portion of the Hugoniot lying between the two phases' curves is a region of mixed phases, where quantities of both phases are present in the shocked material. Table 3.2 lists the pressures at which a number of phase transformations and other shock-induced petrographic features occur during shock compression of rocks and minerals.

The phase transformation produces a kink in the Hugoniot curve similar to the kink caused by yielding at the Hugoniot elastic limit. It also results in a two-shock structure similar to the two-pulse structure found for a plastic wave. In the ideal case, a first shock compresses the material to the pressure at which the Hugoniot curve leaves the phase I curve (Fig. 3.8). The second shock, propagating more slowly than the first, completes the transformation to phase II and raises the material to its final pressure. In actual practice the finite rate of the phase transformation smears out the second shock, so that the sharp onset of the first shock is followed by a gradual rise in pres-

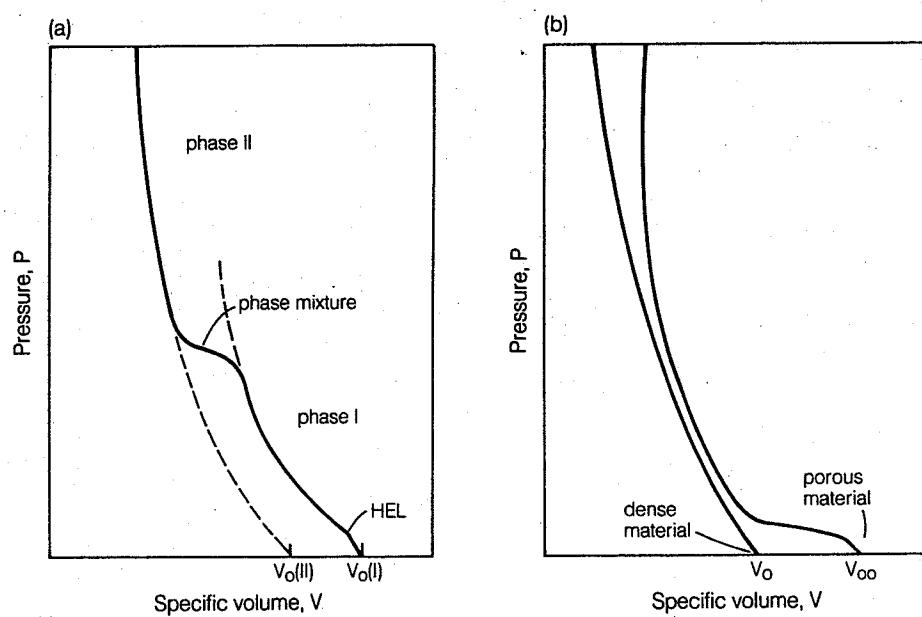


Fig. 3.8 The effect of high-pressure phase transformations (a) and porosity (b) on a material's equation of state in the  $(P, V)$  representation.

TABLE 3.2 Petrographic shock indicators

Material	Indicator	Pressure (GPa)
Tonalite (rock)	Shatter Cones	2-6*
Quartz	Planar elements and fractures	5-35
	Stishovite	15-40
	Coesite	30-50
	Melting	50-65 (?)
Plagioclase	Planar elements	13-30
	Maskelynite	30-45
	Melting	45-65 (?)
Olivine	Planar elements and fractures	5-45 (?)
	Ringwoodite	45
	Recrystallization	45 (?)-65 (?)
	Melting	>70†
Clinopyroxene	Mechanical twinning	5-40 (?)
	Majorite	13.5
	Planar elements	30 (?)-45
	Melting	45 (?)-65 (?)
Graphite	Cubic diamond	13
	Hexagonal diamond	70-140

Unless otherwise noted, data is from Stöffler (1972).

\*Roddy and Davis (1977)

†Jeanloz (1980)

(?) = uncertain

sure. The review by Duvall and Graham (1977) describes the effects of phase transitions in much more detail.

Phase transformations tend to increase the amount of irreversible work done by the shock wave and thus to increase the temperature of the rock after the shock wave has passed. It is easy to see that the increase in specific internal energy, which is given by the area of the right triangle defined by the initial state, the final shock state, and lines parallel to the axes on the  $P$ - $V$  diagram (Fig. 3.10), is larger when phase transformations occur, as shown in Figure 3.8a.

Initial porosity in the target has similar effects. Figure 3.8b shows a Hugoniot curve for a porous material (initial specific volume  $V_{00}$ ) along with the Hugoniot curve for the same material without pores (initial volume  $V_0$ ). The pores collapse and the density of the material increases at relatively low shock pressures (nearly all pores in rocks close below 4 GPa). The Hugoniot curve thus takes a sharp turn to the left, approaching the nonporous Hugoniot.

The porous Hugoniot curve, however, always lies above the nonporous curve because compressed porous material is hotter and thus less dense than nonporous starting material at the same pressure. The internal energy of the compressed porous material is larger than that of the nonporous material because the triangle formed by the Rayleigh line and lines parallel to the  $P$ - $V$  axes is larger (it has a much broader base). Phys-

ically, the porous material has more  $PdV$  work done upon it during compression and so receives more heat. In extreme cases this effect can even reverse the generally negative slope of the Hugoniot curve, causing it to become *positive* on the  $P$ - $V$  plot.

Porosity is not recovered after a shock wave has compressed the material, and thus it contributes strongly to irreversible heating. If it is of sufficiently large magnitude, the presence of initial porosity may delay or inhibit high pressure phase transformations by increasing the temperature of the shock-compressed material. Some experimental investigations, in fact, deliberately employ starting materials with different porosities in an effort to study the effects of internal energy variations on the shock equation of state.

There are many other processes that occur during shock compression of rock materials, not all of which are well understood at the present time. Some of these, such as the formation of shear bands and planar features in minerals, are more important as postshock stress indicators than for their effect on shock propagation. The formation and growth of microcracks during shock compression may have an effect on shock propagation by scattering wave energy, and thus smearing out the wave front, and by subtracting energy from the coherent shock.

Although their mechanisms are poorly understood, a number of shock-induced alterations in rocks and minerals are used as petrologic indicators of peak stress. For instance, peak shock pressures in the range of 30 to 45 GPa transform plagioclase feldspar into an isotropic diaplectic glass known as Maskelynite. A rock that contains Maskelynite has thus been subjected to a shock wave at least this strong.

The most famous shock indicators in quartz-rich rocks are Coesite and Stishovite. These dense phases of silica were first found in nature at Meteor Crater, Arizona, and provided important confirmatory evidence of that crater's impact origin, since volcanic explosions cannot approach the pressures at which Stishovite and Coesite form (greater than 15 and 30 GPa, respectively). Subsequent discovery of Stishovite and Coesite at the Ries Crater in Germany was also accepted as proof of that structure's impact origin. Most recently, the discovery of shocked quartz in the Cretaceous-Tertiary boundary clay in Montana lends increased weight to the hypothesis that the Cretaceous Era was brought to a close by a major impact event.

There is mounting evidence that the course of shock propagation in rocks is highly irregular: petrographic study indicates that shock pressures may vary by a factor of two or more over distances of centimeters or even meters in rock outcroppings that were supposedly uniformly shocked by a large meteoroid impact. It has often

been observed that peak pressures seem to vary on a scale of millimeters in thin sections of shocked rocks. These complexities are too poorly understood to be described here. The interested reader should consult the paper by Grady (1977) for a review of this subject.

Table 3.2 lists a number of petrologic and mineralogic shock pressure indicators, ranging from shatter cones (cone-shaped fractures that commonly occur in shocked bedrock on scales from tens of meters to centimeters and that probably reflect yielding just above the Hugoniot Elastic Limit) to the formation of glass from refractory minerals at shock pressures above about 50 GPa. The topic of shock metamorphism is a large area of study. Readers wishing more information should consult the excellent review by Stöffler (1972).

### 3.4.3 Release from high pressure

The high pressure state induced by an impact is transient. The duration of high pressure is variable, depending upon the size and velocity of the impacting projectile, but ranges from  $10^{-3}$  sec to  $10^{-1}$  sec for 10-m-to 1-km-size projectiles. Although this time is brief, a great many physical changes occur during both shock compression and release.

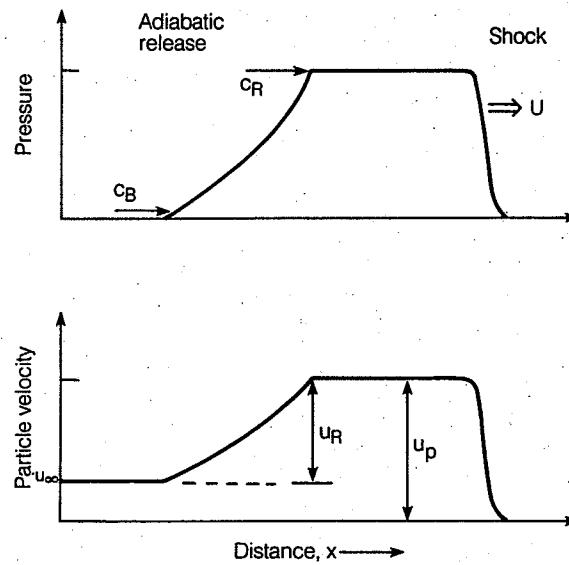
After release from high pressure the decompressed material is seldom in its initial state. Even weak shocks or stress waves leave fractures in brittle solids such as rock. Pore space in soils or sedimentary rocks is crushed out. Stronger shocks deposit a significant amount of heat in the material they compress. Water initially present in pores may vaporize, shattering the rock as it expands. Very strong shocks can deposit enough heat to melt or vaporize the rocks through which they pass.

The high pressure in a shock wave is relieved by the propagation of *rarefaction*, or release, waves from free surfaces into the shocked material. Unlike elastic wave reflection, where the tensional relief wave propagates away from the free surface at the same speed  $c_L$  as the incident compressional pulse, the rarefaction wave from a strong compression generally moves faster than the shock wave. The rarefaction wave's speed  $c_R$  is proportional to the slope of the adiabatic release curve on the  $P$ - $V$  diagram:

$$c_R = \sqrt{\left. \frac{\partial P}{\partial \rho} \right|_S} \quad (3.4.11)$$

where the subscript  $S$  indicates a partial derivative taken at constant entropy.

Although the head of the rarefaction wave travels into the high-pressure region at speed  $c_R$ , the trailing, low-pressure end of the wave moves at a lower speed, on the order of  $c_B$ , if the material that



**Fig. 3.9** Propagation of release waves following a strong shock. The release wave, which is a pressure wave, not a shock, travels at the speed of sound in the shocked material,  $c_R$ . Moving faster than the shock itself, it allows compressed material to expand to low pressure. The strength of the shock declines rapidly when the rarefaction catches up with it. Although the pressure may ultimately drop to, or near, zero, the particle velocity declines to a fixed residual value  $u_\infty$  because of thermodynamic irreversibility in the shock. Porosity, yielding at the HEL, or phase transformations thus tend to increase the residual velocity. This residual velocity is the source of the excavation flow field in impact cratering.

emerges from the shock is still solid, or at the sound speed if vaporization occurs upon unloading. Figure 3.9 illustrates the relation between a shock wave and the rarefaction wave in a one-dimensional geometry.

Material engulfed by the shock wave is accelerated by the sharp pressure gradient at the shock front. Similarly, material affected by rarefaction waves is accelerated down their more gradual pressure gradient. The net velocity of the decompressed material is the vector sum of these accelerations integrated over time. The particle velocity parallel to the direction of propagation of a plane shock,  $u_p$ , can be deduced from the Hugoniot equations (Eqs. 3.4.1 to 3.4.3) coupled with the equation of state (Eq. 3.4.4). Similarly, the change in particle velocity of decompressed material parallel to the direction of propagation of a plane rarefaction wave,  $u_R$ , is deduced from the Riemann invariant:

$$u_R = - \int_{\rho_0}^{\rho} \frac{c_R(\rho) d\rho}{\rho} \quad (3.4.12)$$

The magnitude of this change in velocity is slightly less than that of the particle velocity be-

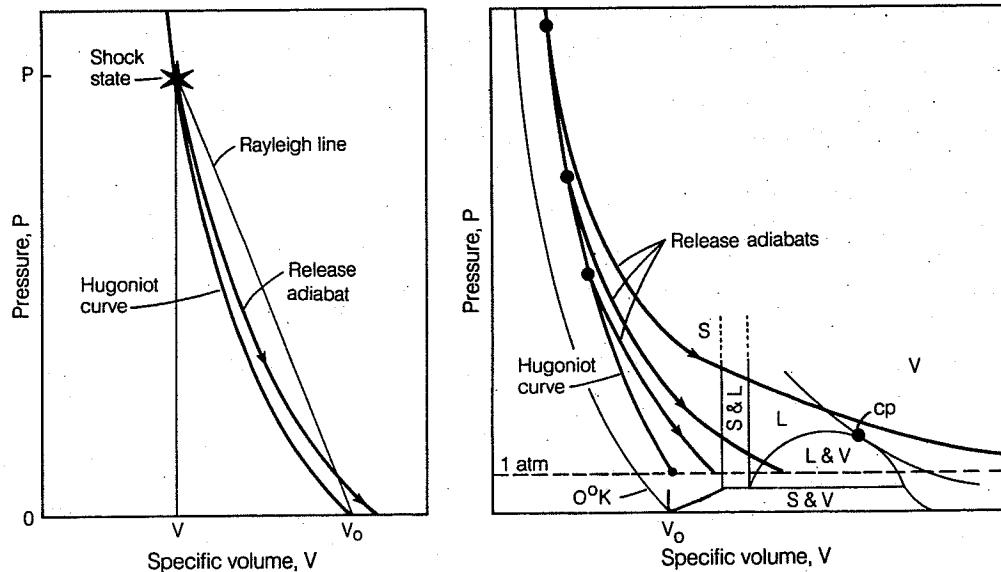
hind the shock (unless the decompressed material vaporizes). The successful application of the velocity doubling rule when applied to strong shocks depends upon the similarity of these two velocities. The particle velocity of material near a free surface from which a shock wave has been reflected is  $u_p + u_R$ . This is nearly equal to  $2u_p$  when  $u_p \approx u_R$ . See the paper by Walsh and Christian (1955). Similarly, the particle velocity in decompressed material behind a propagating shock is the difference between these velocities,  $u_\infty = u_p - u_R$ , which is generally small in comparison to  $u_p$ , but not zero. Thus, material compressed and accelerated by a shock wave is not later decelerated to rest by the rarefaction: a residual velocity  $u_\infty$  remains. This residual velocity plays a central role in impact crater excavation.

The ultimate cause of this difference in particle velocities produced by a shock wave and a rarefaction wave is thermodynamic. The shock wave conserves mass, energy, and momentum as it compresses the material through which it travels. The rarefaction conserves all these and entropy as well. Shock compression is thermodynamically irreversible, whereas rarefaction is reversible and adiabatic. Any process that contributes to the irreversibility of shock compression also enhances the difference between  $u_p$  and  $u_R$ . Yielding at the Hugoniot Elastic Limit, crushing of pore space, or

unreversed phase transformations all act to increase the difference between these two velocities. Since, as will be seen later, the speed of the excavation flow in an impact crater is roughly equal to this difference, it is important to include irreversible phenomena in theoretical attempts to model impact crater excavation. The simplified equations of state used in early modeling efforts generally predicted much smaller craters than observed. These failures were almost certainly due to the inadequate equation of state. Modern efforts, using increasingly realistic equations of state, yield much more accurate crater dimensions.

Release from high pressure, besides accelerating material down the pressure gradient established by release waves, may also result in a change of state. Shock compression deposits a large amount of internal energy in the material. Since a shock is not thermodynamically reversible, much of this energy remains as heat even after decompression. If enough heat remains, the material may be left in the form of a liquid or vapor.

Figure 3.10 illustrates the thermodynamic path of a decompressing material on a  $P$ - $V$  diagram. The left-hand side shows the course of events for a material that returns to a solid (but hotter) state after compression and release. The material has



**Fig. 3.10** Release from shock compression is adiabatic and follows a continuous thermodynamic path on the  $P$ - $V$  plot. The left-hand portion of the figure shows the Rayleigh line connecting the initial and shock states. If the initial pressure  $P_0$  can be neglected, the specific internal energy of the shocked material is equal to the area of the triangle on this plot, by the third Hugoniot equation. Upon decompression, the area between the release adiabat and the Rayleigh line (reckoned positive when the Rayleigh line is above the adiabat, negative when it is below), is the net energy, per unit mass, deposited in the unloaded material. The right-hand portion of the figure schematically illustrates the possible phase changes produced by release from shock compression.  $S$  stands for solid,  $L$  for liquid,  $V$  for vapor. The critical point is labeled  $cp$ . Also shown is the  $0\text{ K}$  equation of state curve for the solid. Since shock increases internal energy, the Hugoniot curve always lies above it.

initial specific volume  $V_0$ . The shock suddenly raises it to pressure  $P$  and specific volume  $V$  on the Hugoniot curve (shown as a star on the curve). The specific internal energy density of the shocked material is equal to the area of the triangle bounded by the Rayleigh line and lines parallel to the  $P$  and  $V$  axes, as shown in Figure 3.10. Following the shock, the material decompresses along an adiabatic curve, also shown. Unlike the Hugoniot, this curve is a thermodynamic path that traces a sequence of states that the material actually passes through (any nonadiabatic processes due to reaction kinetics during release are ignored here, although such complications may occasionally become important). When phase changes do not occur, the release adiabat closely approximates the Hugoniot curve (indeed, in some work the two curves are taken to be identical, although they have very different physical interpretations), except that it decompresses to a slightly larger specific volume than initially. This difference is due to thermal expansion of the hot decompressed material.

Decompression of the shocked material converts much of the internal energy gained during compression into mechanical work. The energy thus converted is equal to the area between the release adiabat (Fig. 3.10) and lines parallel to the  $P$  and  $V$  axes. This is smaller than the area of the triangle bounded by the Rayleigh line. The difference between these areas is equal to the amount of irreversible work done by the shock. Thus, the area between the release adiabat and the Rayleigh line (counted as positive if the release adiabat lies below the Rayleigh line, negative if the reverse) equals the amount of heat appearing in the de-

compressed material. This heat is easy to estimate if the release adiabats approximate the Hugoniot curve. Otherwise, the release adiabats must be determined experimentally.

The right-hand side of Figure 3.10 illustrates the more general case of release to a different state of matter. A generalized phase diagram is projected on the  $P$ - $V$  diagram and several release adiabats originating at different shock pressures are shown. The state of the material after release depends on the peak shock pressure it experiences. At modest shock pressures initially solid material is released as a hot solid. Release from progressively higher shock pressures produces first liquid, then vapor. Note that the Hugoniot curve always lies above the  $P$ - $V$  curve of the solid at 0K because of the shocked material's internal energy content. The release adiabats initially follow the Hugoniot curve closely, diverging only at lower pressures. This behavior is a consequence of energy conservation: the area beneath the release adiabats cannot exceed the area beneath the Rayleigh line.

Melting and vaporization of rocks and minerals are complex phenomena involving the interaction of several phases and large numbers of species. The release adiabats for these substances have not been worked out in detail (either theoretically or experimentally). Only general limits can be placed upon the shock pressures necessary to melt or vaporize a given rock or mineral. This pressure is, moreover, a function of the initial porosity of the starting material, more porous substances requiring less shock to melt or vaporize than dense substances. Nevertheless, Table 3.3 indicates approximately the shock pressures needed to induce

TABLE 3.3 Shock induced melting and vaporization

Material	Pressure in GPa for:				
	Melting		Vaporization		Source
	Onset	Complete	Onset	Complete	
Water ice (-50°C)	3	10	26‡	≥100	Kieffer and Simonds (1980)
Water	—	—	5	~100	
Halite	35	?	65	280	Allen (1967)
Limestone*	66	?	110	370	Allen (1967)
Dolomite	84	?	130	430	Allen (1967)
Dry Tuff	43	?	74	260	Allen (1967)
Wet Tuff	54	?	90	320	Allen (1967)
Oil Shale	73	?	120	400	Allen (1967)
Granite	78	?	130	520	Allen (1967)
Gabbroic Anorthosite†					
lpp	97	110	190	790	Ahrens and
hpp	43	52	102	590	O'Keefe,
Iron	220	260	420	1680	(1977)

\*Calcite undergoes decarbonation,  $\text{CaCO}_3 \rightarrow \text{CaO} + \text{CO}_2$ , beginning at 45 GPa and reaching completion by 70 GPa (Kieffer and Simonds, 1980).

†Gabbroic Anorthosite transforms from a low-pressure-phase assemblage (lpp) to a high-pressure-phase assemblage (hpp) at 15 GPa. The lpp values listed are theoretical extrapolations, supposing that no phase transformation occurs.

‡The "onset of vaporization" for water and water ice depends upon the partial pressure of the environment into which the shocked material expands. The tabulated pressure is the pressure from which the release adiabat passes through the critical point.

phase transitions in dense geologic materials. For further information see the paper by Ahrens and O'Keefe (1972).

## FURTHER READING

### General References

The best work by far on shock waves is the two-volume book:

Ya. B. Zel'dovich and Yu. P. Raizer, *The Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, Academic Press, New York, 1967.

Chapter I gives a clear and complete discussion of shock waves in general, while chapter XI is devoted to shock waves in solids. Two other useful reviews are

L. V. Al'tshuler, Use of shock waves in high-pressure physics, *Soviet Physics USPEKHI* 8: 52-89 (1965)

and

M. H. Rice, R. G. McQueen, and J. M. Walsh, Compression of solids by strong shock waves, *Solid State Physics* 10: 1-63 (1958).

### Special Topics

A clear and complete exposition of elastic wave propagation, reflection, and fracture that requires little mathematical sophistication can be found in

J. S. Rinehart, On fractures caused by explosions and impacts, *Quat. J. Colo. Sch. Mines* 55: 1-155 (1960).

The most detailed discussion of yielding at the Hugoniot Elastic Limit is in

G. R. Fowles, Shock wave compression of hardened and annealed 2024 Aluminum, *J. App. Phys.* 32: 1475-1487 (1961).

The standard reference on phase transformations during shock compression is

G. E. Duvall and R. A. Graham, Phase tran-

sitions under shock-wave loading, *Rev. Mod. Phys.* 49: 523-579 (1977).

The effects of porosity are best described by Zel'dovich and Raizer and are also treated by Al'tshuler, both cited above. The most complete work on the petrographic effects of shock is

D. Stöffler, Deformation and transformation of rock-forming minerals by natural and experimental shock processes, *Fortschritte der Mineralogie* 49, 50-113 (1972). In English.

A good review of the rapidly advancing understanding of the complex processes that occur behind a shock front is given by

D. Grady, Processes occurring in shock wave compression of rocks and minerals, in M. H. Manghnani and S.-I. Akimoto (Eds), *High Pressure Research: Applications in Geophysics*, Academic Press, New York, pp. 389-438 (1977).

The velocity doubling rule and the thermodynamics of release without phase change are well presented in

J. M. Walsh and R. H. Christian, Equations of state of metals from shock wave measurements, *Phys. Rev.* 97: 1544-1556 (1955).

The best discussion of the thermodynamics of release from shock compression with phase changes is by

T. J. Ahrens and J. D. O'Keefe, Shock melting and vaporization of lunar rocks and minerals, *The Moon*, 4: 214-249 (1972).

This last work also contains a great deal of data and estimates of the melting and vaporization conditions for many geologic materials.

A discussion of the effects of volatiles included in the pore space of target rocks during release from high pressure can be found in

S. W. Kieffer and C. H. Simonds, The role of volatiles and lithology in the impact cratering process, *Rev. Geophys. Space Phys.* 18: 143-181 (1980).

## IV. CRATERING MECHANICS: CONTACT AND COMPRESSION STAGE

### 4.1 THE THREE STAGES OF IMPACT CRATERING

The impact of a meteorite moving at tens of kilometers per second on a planetary surface initiates an orderly—but rapid—sequence of events that produces the final crater. The formation of an impact crater is a continuous process that begins when the meteorite (the “projectile”) first contacts a planetary surface (the “target”) and ends with the final motions of debris around the crater. However, it is convenient to divide this process into several stages, each of which is dominated by a different set of physical phenomena.

These stages grade into one another and it is not possible to define precisely when one ends and the next begins. An artificial division into stages might thus seem pointless, since it obscures the fundamental continuity of the cratering process. However, the real purpose of the division is to enhance the ability of the human mind to comprehend the vast array of processes that occur during an impact. A full description of a cratering event is possible through numerical modeling techniques, but these do not produce readily grasped generalizations. Numerical modeling provides useful quantitative data for particular impact events but it does not lead to an intuitive understanding of the process itself. Such understanding is gained only by grouping similar phenomena together and dividing the process into stages. First proposed by Gault et al. (1968), three stages are generally recognized. Each has several different aspects and may thus be further subdivided. These stages are contact and compression, excavation, and modification.

#### 4.1.1 Contact and Compression

The first stage of impact cratering begins when the projectile contacts the target surface. Unless the target planet has an atmosphere (a case treated in Chapter 11), the approaching projectile does not affect the target in any way until it actually strikes. The swiftly moving projectile then pushes target material out of its path, compressing it and accelerating it to a large fraction of the impact velocity. At the same time, the target’s resistance to penetration decelerates the projectile. These velocity changes are mediated by shock waves: material in

the contact zone between the projectile and target is strongly compressed, creating shock waves at the boundary between the compressed and un-compressed material.

These shock waves originate at the point (or points, for irregular impactors) where the projectile first touches the target’s surface. As the projectile continues its plunge into the surface, the shock fronts spread and propagate into both projectile and target. Shock pressures developed during the early stages of most hypervelocity impacts generally reach hundreds of GPa and thus far exceed the yield strength of both projectile and target. Both materials may either melt or vaporize upon unloading from such pressures. High-speed “jets” of strongly shocked material squirt out from the interface between the projectile and target. Most of the projectile’s initial kinetic energy is transferred to the target: the underlying rocks are compressed, heated, and accelerated to high speed.

The contact and compression stage ends after the projectile has unloaded from high pressure. The principal result of this stage is the transfer of most of the projectile’s kinetic energy to the target. This stage is characterized by high pressures, temperatures, and velocities, all localized within a region comparable in size to the projectile. Contact and compression is the shortest of all the stages, lasting only as long as the shock wave and subsequent rarefaction take to traverse the projectile. Its duration depends upon the projectile’s size, composition, and impact velocity, ranging from ca.  $10^{-3}$  seconds for a 10-m diameter silicate projectile striking at 10 km/second to ca.  $10^{-1}$  seconds for a 1 km diameter projectile under similar conditions. For most impacts, therefore, contact and compression is over in literally less than the “wink of an eye.” This stage lasts a second or more only for the very largest impacts.

#### 4.1.2 Excavation

The excavation stage begins immediately after the contact and compression stage ends. During this next stage a more-or-less hemispherical shock wave propagates into the target, weakening as it expands and engulfs more material. This shock wave and the following rarefaction set target material in motion, initiating a subsonic excavation

flow that eventually opens the crater. The resulting crater's diameter is many times larger than the projectile that produced it (this is one of the major differences between high velocity impacts and the more familiar low velocity penetration-type craters made, for example, by falling rocks or military shells). The ejecta curtain forms and begins to blanket the terrain adjacent to the final crater during this stage. The target's material strength and gravity become important near the end of excavation.

The excavation stage lasts many times longer than the contact and compression stage, requiring seconds or minutes to reach completion, depending upon the crater size. Chapter 5 treats this stage in detail.

#### 4.1.3 Modification

The modification stage begins after the crater has been fully excavated. Excavation produces a bowl-shaped "transient crater" that generally collapses under gravity. During modification, loose debris slides down the steep interior walls of small craters, pooling on the floor of the final bowl-shaped depression. Large craters collapse more spectacularly: slump terraces form on the walls

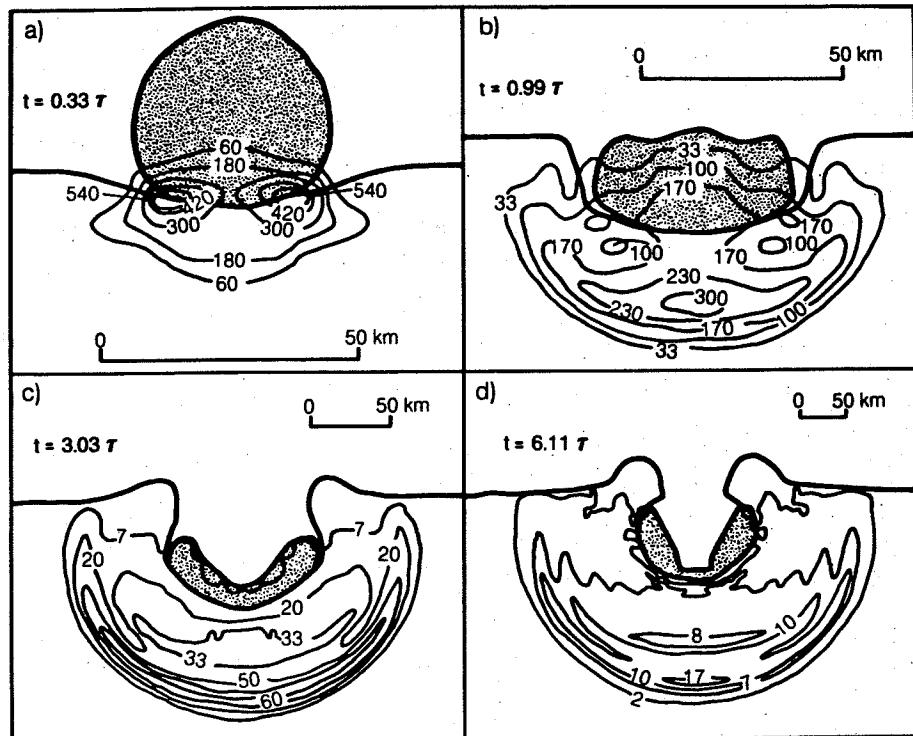
and central peaks rise in the interior. Great mountain rings may appear in and around still larger craters if certain conditions in the target are met. On a much longer timescale, isostatic rebound may follow collapse, eventually flattening the crater to a mere albedo feature. These topics are treated in Chapters 8 and 9.

The remaining sections of this chapter are an in-depth discussion of the contact and compression stage and the processes that dominate this phase of impact cratering. A thorough understanding of shock waves, at the level of Chapter 3, is a prerequisite for complete comprehension of this material.

#### 4.2 VERTICAL IMPACT OF A SPHERICAL PROJECTILE

The contact and compression stage of impact is most easily visualized in a simplified geometry: a spherical projectile impacts vertically on a homogeneous target bounded by a plane free surface. This geometry is also the easiest to treat by numerical methods because its axial symmetry reduces the dimensionality of the problem from three to two.

Figures 4.1a to 4.1d illustrate the events during



**Fig. 4.1** Four snapshots are shown of the vertical impact of a 46.4-km diameter iron projectile on a gabbroic anorthosite target at 15 km/second. The first three frames illustrate different phases in the contact and compression stage and the last frame is a very early phase of the excavation stage. The contour values are pressures in GPa. Times shown are in units of  $\tau$ , Equation 4.2.1. See the text for individual discussion of the frames. Do not overlook the changes in scale from one frame to the next. After O'Keefe and Ahrens (1975).

the contact and compression stage for this kind of impact. These figures are derived from a particular numerical computation (the impact of an iron projectile on a gabbroic anorthosite target at 15 km/second), but they incorporate the features common to all high velocity impacts of the vertical type. Special features, such as the presence of layering or volatiles in the target are not shown here. The effects of such circumstances are largely unknown at present. The understanding summarized in these figures is, however, adequate to explain all that is presently known about this type of impact, either experimentally or observationally.

Figure 4.1a shows the projectile just after it has contacted the target surface. A small region of high pressure develops along the interface. This region encompasses both shocked target and shocked projectile. Both target and projectile are compressed and begin to distort in shape. The rear of the projectile continues its plunge into the target at unabated velocity, not yet affected by the events occurring at the leading edge. Target material outside of the shocked zone is similarly undisturbed until the actual arrival of the shock front.

The figure shows two small "islands" (actually sections of a torus in three dimensions) of extra-high pressure centered on the contact between the projectile and target. These high pressure regions are associated with the process of jetting, discussed further in Section 4.5. They result from the oblique convergence of the curving sides of the projectile and the initially flat target surface. A hot, heavily shocked mixture of projectile and target squirts or "jets" from these high pressure zones at speeds several times faster than the projectile itself. These jets involve relatively minor amounts of material in vertical impacts, but may be more important in oblique impacts. Such jets have been observed in the laboratory as low-angle streams of incandescent melt or vapor.

The shock wave propagating into the projectile eventually reaches its rear surface. Figure 4.1b illustrates this event. At this time the projectile-target interface has been pushed roughly one-half the projectile diameter into the target. The projectile is completely engulfed by the shock wave and its entire mass is compressed and decelerated to a fraction of the initial impact velocity. An increasing amount of the target is engulfed by the shock and correspondingly accelerated to high speed. Shock pressures greatly exceed material strength, which can therefore be entirely neglected at this stage, and the projectile is becoming distorted as it flows hydrodynamically.

The shock wave in the projectile is reflected as a rarefaction or release wave after it has traversed the projectile and reached its rear surface. The rarefaction travels at the speed of sound in the compressed material, unloading it to near-zero

pressure and further decelerating it, as described in Section 3.4.3. The unloaded projectile may undergo phase transformations, either melting or vaporizing if its initial shock pressure is sufficiently high. Figure 4.1c illustrates this last phase of the contact and compression stage. The shock wave in the target continues to propagate outward and begins to approximate a hemisphere centered roughly one projectile diameter below the surface. The projectile is greatly distorted and now lines the growing crater cavity. Most of the projectile material is vaporized at this stage, although the vapor has not yet begun to expand.

Figure 4.1d illustrates events at the beginning of the excavation stage, after the end of the contact and compression stage. Note the basic continuity underlying the artificially distinguished "stages." The projectile and part of the target have unloaded to near-zero pressure, and the unloaded material is moving rapidly away from the impact site. The projectile still lines the inside of the expanding crater cavity, which is hemispherical by this time. Some of the target material has undergone phase transformations. Because of the downward propagation of the rarefaction wave, the shock wave has become an isolated pressure high that fills an expanding hemispherical shell. The ejecta curtain begins to form as the rim of the crater cavity moves upward and outward. The vaporized projectile is expanding into the crater cavity from its previous position on the cavity walls.

The net result of the contact and compression stage is to convert the kinetic energy of the projectile into internal energy of both projectile and target and to produce a strong, roughly hemispherical, shock wave propagating into the target. All of these energy conversions take place in a region roughly the same size as the projectile and over a time interval  $\tau$  comparable to the time required for the projectile to traverse a distance equal to its own diameter  $L$ . The duration of the contact stage is thus roughly

$$\tau = L/v_i \quad (4.2.1)$$

where  $v_i$  is the initial velocity of the projectile.

For impact velocities exceeding a few kilometers per second, the average energy per unit volume during this stage is similar to (or even larger than) the energy per unit volume of chemical explosives. This surprisingly high energy density initiated the modern understanding of impact cratering by suggesting to investigators of the 1920s and 1930s that the process of crater excavation is similar to an explosion. The expansion of this concentrated energy source is largely independent of the details of how it is established, producing many similarities between the results of explosions and impacts. This impact-explosion analogy is discussed in more detail in Chapter 7.

The near-independence of the final crater's shape, size, etc., from the details of the energy

source means that parameters such as the impactor's velocity, mass, or even angle of impact (within broad limits) cannot be separately determined from the final crater. This is either a blessing or a hindrance, depending upon whether one is interested in the final crater or the initial projectile. In particular, this independence means that all but the most oblique impacts produce circular craters. This fact answered the most cogent objection nineteenth-century astronomers had to the impact origin of lunar craters (see Chapter 1). The energy coupling between the projectile and target in the contact and compression stage is, however, strongly dependent upon the obliquity of the impact. The next section describes what little is known about the earliest stages of an oblique impact.

### 4.3 OBLIQUE IMPACT

No meteorite strikes a planetary surface exactly vertically. Although some impacts may be so nearly vertical that the difference is negligible, the most probable angle of impact of a randomly incident projectile is 45 degrees. The probability of a meteorite's impacting a given surface area at an angle between  $\theta$  and  $\theta + d\theta$  is proportional to  $\sin \theta \cos \theta d\theta$  independent of the target planet's gravitational field. For a proof of this remarkable fact, see Shoemaker (1962). This probability is zero for vertical or grazing impacts ( $\theta = 90^\circ$  or  $0^\circ$ ) and reaches a maximum at  $\theta = 45$  degrees. Oblique incidence of the projectile is thus an important facet of impact cratering.

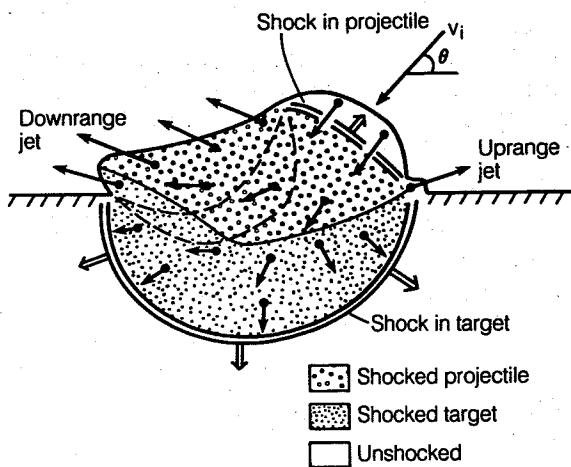
Although all but the most oblique hypervelocity impacts (impact angle less than 10°) produce circular craters (see Sec. 5.6.1), their ejecta blankets have a bilaterally symmetric "butterfly wing" pattern that becomes progressively more marked as the obliquity increases (see Sec. 6.3.6). Phenomena such as ricochet of the projectile, elongated craters, and downrange streaks appear at very low angles of incidence (less than 6° at impact speeds of 6 km/second or more). The initial coupling of energy into the target is less efficient for oblique impacts and the resulting craters are smaller for a given impactor mass and velocity (see Sec. 7.9).

Unfortunately, oblique impacts have not yet been studied in much detail. There are a number of reasons for this omission. First, few experimental facilities are capable of varying the angle of impact. The nearly unique tilting vertical gun facility at NASA's Ames Research Center has produced most of the data that currently exist on small-scale oblique impacts. Second, oblique impacts are inherently three-dimensional and, until recently, were beyond the reach of numerical modeling. This situation is changing rapidly as computers increase in both memory capacity and speed. Finally, it has long been known that the

angle of impact has little effect on the shape of the final crater, except for very low angles of incidence. This seems to have discouraged any real efforts to attain more than a qualitative understanding of oblique impacts.

Events during the compression stage, however, are very sensitive to the obliquity of the impact. Ricochet of the projectile, if it occurs, begins during this stage. Figure 4.2 represents the early phases of the contact and compression stage for an oblique impact. It is drawn at a time comparable to that of Figure 4.1a, when the projectile is not yet completely engulfed by the shock. Figure 4.2 is generalized from several exploratory three-dimensional numerical studies and a "2½-dimensional calculation (oblique incidence of a long rod sidewise onto a flat plate, which is a two-dimensional, plane strain, problem that incorporates some aspects of a fully three-dimensional oblique impact).

The principal difference between vertical and oblique impacts, shown in Figure 4.2, is the fate of the projectile. Projectile material is first compressed by a shock that originates at the projectile-target interface and then propagates into the projectile. The projectile's vertical velocity component is decreased by the shock, but its horizontal component is still large. Rarefactions from the projectile's free surface decrease its vertical velocity still further, and may even result in an upward component, as illustrated in Figure 4.2. The



**Fig. 4.2** Schematic illustration of the early compression and jetting phase in a moderately oblique impact. Shocks form in both projectile and target. A narrow, high-velocity jet spouts from the rear of the projectile. Rarefactions (dashed lines) unload the front of the projectile even while the rear is entering the shock. The projectile and some target material form a down-range jet. In spite of the obliquity of the impact, a crudely hemispherical shock propagates into the target. This shock is not as strong as it would be in a vertical impact at the same velocity.

shocked and unloaded projectile material thus forms a plume that may proceed upward above the extrapolated ground track of the original projectile.

The shock wave in the target is similar to that produced by a vertical impact, but the target material has a horizontal velocity component perpendicular to the line of approach that is probably responsible for the asymmetric ejecta patterns discussed in Section 6.3.6. The projectile penetrates less deeply in an oblique impact than in a vertical one and deposits a smaller fraction of its energy in the target.

Elementary geometry suggests that the energy deposited in the target during the contact and compression stage fills a canoe-shaped elongated trough in the target's surface, rather than the hemispherical region created by vertical impacts. The width of this trough is roughly equal to the projectile diameter  $L$  and its length is roughly  $L/\sin \theta$ . If the impact is extremely oblique, this trough is much longer than it is wide and the projectile will create a nearly cylindrical shock wave centered on the projectile's trajectory, accelerating material sideways away from its track as it plows a furrow into the target.

The final crater will be circular if its diameter  $D$  is much larger than the trough length  $L/\sin \theta$ , because the shock wave produced during contact and compression spreads enough during excavation to obscure the details of its origin. However, if the trough length is comparable to the final crater's diameter, the crater will be elliptical rather than circular.

The duration of contact and compression for oblique impacts is longer than for vertical impacts. As before, a rough measure of the duration is the time it takes for the projectile to travel a distance equal to its own diameter. However, for oblique impacts the appropriate velocity is the *vertical component* of the impact velocity,  $v_i \sin \theta$ . Thus, in analogy to Equation 4.2.1, the duration  $\tau$  of the contact and compression stage for an oblique impact is roughly

$$\tau = L/(v_i \sin \theta). \quad (4.3.1)$$

Equation 4.2.1 is a special case of this equation for  $\theta = 90^\circ$ .

Figure 4.3 illustrates the results of three "2½-D" computations of the early phases of an oblique impact. The material equations of state and impact velocity are the same in each of the three computations. The angles of impact are, from the top down, 45 degrees, 25 degrees, and 10 degrees. All are shown at about the same time,  $t = 1.8L/v_i$ . The timescale's dependence on the angle in Equation 4.3.1 is qualitatively verified by these figures, since contact and compression is nearly complete in the top (steepest angle) frame, still underway in the middle (intermediate angle)

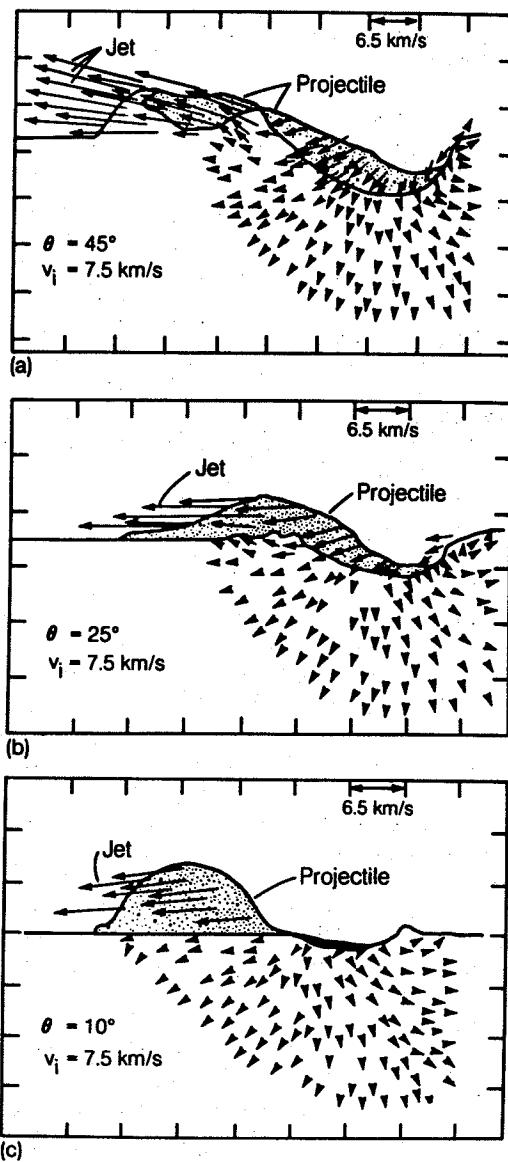


Fig. 4.3 Early time velocity fields of oblique impacts in a series of "2½-D" numerical computations. A silicate impactor moving at 7.5 km/second strikes a plane silicate target at 45, 25, and 10 degrees. Although a vertical impact at this velocity is only capable of melting the silicates, jetting conditions are satisfied in the 45- and 25-degree impacts, and both projectile and target material is vaporized. The 10-degree angle is too shallow for jetting and the projectile remains solid, simply skidding across the target surface. All are at nearly the same scaled time,  $t = 1.8L/v_i$ . After O'Keefe and Ahrens (1985).

frame, and just beginning in the bottom (lowest angle) frame. In each case the projectile deforms into a downrange plume, while the velocity in the target clearly decreases as the angle of impact decreases, which implies that the final crater's size

will be smallest for the most oblique impact. Another interesting feature is the small region of downrange flow *behind* the projectile. The downrange plume is composed mainly of vaporized projectile material in the first two cases, the vaporization at this relatively low velocity of 7.5 km/second being due to the pressure multiplication that occurs in jetting (Sec. 4.4). This pressure multiplication takes place within a small region adjacent to the projectile-target contact similar to the islands of high pressure in Figure 4.1a. Vaporization does not occur for the lowest ( $10^\circ$ ) angle of incidence at this velocity.

Although the 2½-D flow fields illustrated in Figure 4.3 possess many features expected in full three-dimensional flows, they cannot be linked to the most characteristic feature of oblique impacts: the bilaterally symmetric ejecta blanket. Furthermore, 2½-D calculations are known to overestimate shock pressures by about a factor of two. Computations like these are useful, but preliminary, efforts to understand the presently obscure early stages of oblique cratering.

Little further can be said on this important but unexplored subject. The remainder of this chapter is a more-or-less standard exposition of the phenomena occurring during the contact and compression stage. Most of this work is distorted by the tacit assumption that the impact is vertical. It is important to keep this in mind when applying these results to real impact craters.

#### 4.4 JETTING IN IMPACTS

The highest speed ejecta are thrown out during the earliest stage of impact cratering, when the projectile first contacts the target. The oblique convergence of portions of the projectile with the target's surface generates the high pressure "islands" seen in Figure 4.1a, from which jets of highly shocked, melted or vaporized, material squirt into the adjacent open space at speeds usually faster than the projectile itself.

Impact-induced jetting was first observed by D. E. Gault during hypervelocity impact experiments in the early 1960s. High-framing-rate cameras revealed an early-time spray of incandescent debris shooting rapidly away from the impact site at a low angle to the target's surface. The spray was evidently molten material even at impact velocities of 6 km/second, which would not normally have been expected to melt either target or projectile. Higher impact velocities can be expected to eject superheated vapor, which may even be strongly ionized. Most of these early experiments involved vertical impacts, where jetting, while spectacular, is not quantitatively important.

The jetting phenomenon is well understood in the context of armor-penetrating shaped charges,

but it has received little attention from the viewpoint of impact cratering because it seemed to involve such a small fraction (less than about 10 percent) of the projectile's mass or energy. Recent work, however, suggests that jetting may play a major role in oblique impacts and that its neglect is therefore unwarranted.

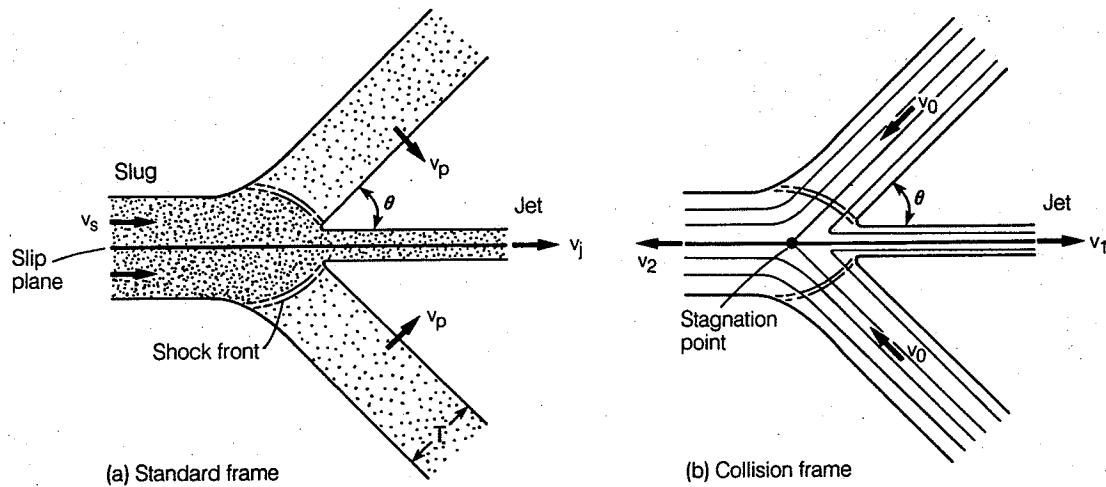
Unlike the well-analyzed collapse of the liner of a shaped charge, however, jetting in impacts cannot be approximated by a steady, symmetric, convergence of identical thin plates at a constant angle. The convergence angle and cross-sectional area of the colliding objects vary rapidly during an impact and it is unlikely that a steady state is even approximately established. Unfortunately, there is little information on the jetting process in these circumstances. The 2½-D oblique impact computations illustrated in Figure 4.3 are some of the few investigations that show jet formation, but they have not yet been systematically analyzed from the viewpoint of jetting.

The following section is a brief presentation of the current knowledge about jetting from, necessarily, the steady convergence point of view. It will have to serve as qualitative introduction to jetting until a more comprehensive understanding of the role of jetting in impact processes is attained. This section is somewhat more mathematically demanding than the rest of the book. I have included it because I think that jetting is more important than most authors on impact have supposed and because there is no comparable presentation in the literature. Some readers may wish to skip to Section 4.4.2, as the material in the next section is not required for understanding the remainder of the book.

##### 4.4.1. Jetting from the collision of thin plates

Jetting was originally studied during World War II, when it constituted the secret behind armor-penetrating weapons. Workers at that time treated it as a hydrodynamic process, realizing that when two thin fluid sheets collide at a small angle  $2\theta$ , a thin high-speed jet emerges along the bisector (see Figure 4.4). The jet's speed tends toward infinity as the angle tends toward zero. Since momentum must be conserved in the collision, the amount of material in the jet diminishes toward zero as the angle vanishes. A much lower velocity "slug" is composed of the fluid that does not enter the jet.

Later work revealed that the process is not entirely hydrodynamic. The formation of shock waves in the plates suppresses jet formation until the convergence half-angle exceeds some minimum  $\theta_{cr}$ . This angle is a complex function of the plates' equation of state and the convergence velocity. In general, however,  $\theta_{cr}$  is small (approaching zero) at low impact velocities but rises monotonically with increasing velocity until at 5 to 10



**Fig. 4.4.** Jetting in the collision of thin plates. Figure 4.4a illustrates the configuration of shocked (heavy stipple) and unshocked (light stipple) material as it passes through a pair of detached shocks in the "standard frame." Figure 4.4b shows the stationary stagnation point and streamlines in the "collision frame." The two frames differ by a velocity transformation of  $v_p \csc \theta$  along the horizontal axis. Both diagrams are drawn for symmetric jetting. A velocity differential occurs across the slip plane in asymmetric jetting (see text).

km/second it attains nearly constant values, ranging between 15 and 20 degrees, depending upon material. The interested reader will find more details in the references cited at the end of this chapter.

Figure 4.4 illustrates the major features of jetting. The two plates in Figure 4.4a each move at velocity  $v_p$  perpendicular to their faces and strike each other at half-angle  $\theta$ . This coordinate frame is here called the "standard frame" since most publications on jetting cite the velocity  $v_p$ . Many authors call this the "laboratory frame"; however, since the coordinate frame moves with respect to the target in an impact, this usage is avoided here. The plates approach each other at a net velocity equal to the sum of the components of  $v_p$ , perpendicular to the axis of symmetry,  $2v_p \cos \theta$ .

The "standard" coordinate frame of Figure 4.4a does not always exist. Although it is always possible to transform to a coordinate frame where one of the objects in a collision has its velocity vector perpendicular to its surface, the second object will not generally meet the same condition. When it does, it is called "symmetric jetting." Otherwise, it is "asymmetric jetting." This latter case, by far the most common in impact cratering, has not been well studied. The usual expedient is to introduce a slip plane (marked in Figure 4.4a) across which the velocity changes discontinuously. A coordinate frame in which the upper and lower plates have a mean velocity of  $v_p \sin \theta$  along the slip plane preserves the results on jet velocities presented below.

The material of the upper and lower plates slides differentially along the slip plane, one moving a little faster than the mean velocity, the other moving an equal amount slower. If the two plates

differ in material properties, a slip plane is also needed, even if the velocity vectors apparently satisfy the conditions for a symmetric jet. A difference of the plates' velocity components out of the plane of the figure must likewise be treated by the slip plane prescription. In reality, a velocity discontinuity across the slip plane is not stable and turbulent mixing of material from either plate will occur across it. Such mixing does not, however, affect the major results on jet mass and velocity since these depend only upon momentum conservation, which is valid in any case.

Figure 4.4b illustrates the other commonly used coordinate frame, the "collision frame." This frame is related to the standard frame of Figure 4.4a by a velocity translation along the symmetry axis. Its principal virtue is that the flow is steady in this frame (for symmetric jetting), so that the stagnation point is stationary and streamlines can be defined. The two plates' velocity vectors  $v_0$  are parallel to their faces in the collision frame. Vector analysis shows that these velocities are related by  $v_0 = v_p \cot \theta$ . The jet velocity in the collision frame,  $v_1$ , is generally found to be a fraction  $f$  of the plate velocity  $v_0$ :  $f = v_1/v_0$ . This fraction  $f$  is rigorously equal to 1 in the hydrodynamic theory and only departs slightly from unity with realistic equations of state. The jet velocity in the standard frame,  $v_j$ , is derived from  $v_1$  by adding the velocity difference  $v_p \csc \theta$  between the standard frame and the collision frame to the jet velocity  $v_1 = v_p f \csc \theta$  in the collision frame, obtaining

$$v_j = v_p (\csc \theta + f \cot \theta) \quad (4.4.1)$$

The jet velocity increases as the angle  $\theta$  decreases. The maximum jet velocity occurs at the critical angle,  $\theta = \theta_{cr}$ .

The maximum pressure in the jetted material is estimated from Bernoulli's equation, which states that the sum of the specific enthalpy,  $H = E + PV$ , and the specific kinetic energy,  $\frac{1}{2}v^2$ , is a constant along streamlines in steady flow. The specific enthalpy is thus highest at the stagnation point (Fig. 4.4b), where the velocity is zero. Material in the jet is compressed to slightly lower pressures, but the stagnation pressure will serve as a guide. Using the third Hugoniot equation (3.4.3) and neglecting any initial pressure  $P_0$  and enthalpy  $H_0$ , an implicit equation for the pressure  $P_s$  at the stagnation point is obtained:

$$v_0^2 = P_s [V_0 + V(P_s)] \quad (4.4.2)$$

This is not a simple algebraic equation because the right-hand side involves both pressure  $P_s$  and specific volume  $V(P_s)$ . These are linked by the equation of state (3.4.4) so that a solution is possible. If the material has a linear shock velocity-particle-velocity equation of state (3.4.9), then Equation 4.4.2 can be reduced to a quadratic equation.

The pressures computed from Equation 4.4.2 are largest near the critical angle  $\theta_{cr}$  because  $v_0$  is then at its maximum for a given "impact" velocity  $2v_p$  (in actuality  $v_p$  must itself be computed from the impactor velocity  $v_i$  and the geometry of the impact, but in most geometries  $v_p$  and  $v_i$  are close). This pressure often exceeds that in a head-on collision ( $\theta = 0$ ) by a factor of two or more. At high impact velocity, the factor may exceed five. The pressure multiplication for oblique convergence explains the islands of high pressure between the projectile and target in Figure 4.1a.

This pressure multiplication may cause melting or vaporization at impact velocities that would not normally be capable of producing changes of state. Thus, silicate melt may appear in jets during 1 to 2 km/second impacts, although it is not expected in quantity until the impact velocity exceeds about 5 km/second. Jet-produced phase changes may be particularly important in oblique impacts or for the collision of irregular bodies where the geometry favors oblique convergence of their surfaces. Melt attributed to jetting has even been identified *inside* shocked rocks, where oblique convergence of the interior surfaces of collapsing pores causes jetting into the remaining open spaces.

The amount of material ejected by the jet can be estimated from momentum conservation in the hydrodynamic model. This gives the mass per unit length of the jet  $m_j$  in terms of the mass per unit length of the plates,  $m$ , as  $m_j = m/(1 - \cos \theta)$ . However, this expression does not work well for  $\theta$  in the vicinity of the critical angle. A better estimate is an empirical relation (deduced from numerical computations) for the width  $w$  of the narrowest part of the jet,

$$w = T(\theta - \theta_{cr})/(\pi - \theta_{cr}) \quad (4.4.3)$$

where  $T$  is the thickness of the plate. In this expression, the jet width vanishes at the critical angle  $\theta_{cr}$  and equals the thickness of one plate  $T$  when  $\theta = \pi$  (radians). The mass flux in the jet is obtained by multiplying  $w$  times the jet velocity  $v_i$  in the collision frame times the jet density, best approximated by the uncompressed density. The mass per unit length of the jet is  $m_j = \rho_0 w$ , per unit of distance perpendicular to the plane of the figure.

#### 4.4.2 Duration of jetting

Jetting is a very early time phenomenon; it is usually complete before the projectile is fully compressed. For a spherical projectile it is over when the projectile has plunged halfway into the target, at a time  $\tau_j \approx L/2v_p$ . If the projectile has other irregularities of scale  $L_i$ , then the jetting associated with these irregularities lasts only  $L_i/v_p$ . These durations are, of course, longer for oblique impacts.

In spite of jetting's short timescale, the jet velocities are so high that a significant amount of material may be ejected in the jet, given favorable geometry. The total jetted mass, composed of comparable quantities of both target and projectile material, theoretically approaches 80 percent of the projectile mass for a head-on collision of equal diameter spheres. Oblique impacts may similarly produce jets comparable in mass to the projectile. Thus, although jetting has been largely ignored in most of the past work on impact cratering, the recent interest in oblique impacts is bringing an awareness that jetting may play an important role in the impact process.

#### 4.5 SHOCK PRESSURES DURING CONTACT AND COMPRESSION

The largest shock pressures attained in an impact event occur during the contact and compression stage. These pressures far exceed those generated by volcanic or chemical explosions, but are less than the maximum pressures produced by nuclear detonations—at least for impact velocities less than 100 km/second. Table 4.1 lists the maximum shock pressure attained for a number of projectile-target pairs over a range of impact velocities. These pressures range between 100 and 1000 GPa for most situations of geologic interest. The corresponding temperatures reach tens of thousands of degrees and vaporized material may be significantly ionized.

The very highest pressures during this high-pressure stage are attained locally, where the projectile and target's surfaces converge obliquely and the process of jetting, discussed in the last section, dominates. The mean pressures, however, are somewhat lower, although still high by conventional standards. Figure 4.1b illustrates that, after the local pressure highs of the early contact

TABLE 4.1 Maximum shock pressures in vertical impacts

Impact Velocity (km/second)	Shock Pressure, GPa, for:		
	Iron on Gabbroic Anorthosite*	Gabbroic Anorthosite on Gabbroic Anorthosite*	Gabbroic Anorthosite on Ice†
5.0	82.5	62.0	29.2
7.5	153.0	99.0	57.1
15.0	481.0	304.0	185.0
30.0	1590.0	1010.0	633.0
45.0	3360.0	2130.0	1330.0

\*After Ahrens and O'Keefe (1977), by the planar impact approximation.

†Computed from Equation 4.5.4 using  $C = 7.71$  km/second,  $S = 1.05$ ,  $\rho_0 = 3965$  kg/m<sup>3</sup> for hpp Gabbroic anorthosite and  $C = 1.317$  km/second,  $S = 1.526$ ,  $\rho_0 = 920$  kg/m<sup>3</sup> for ice (see Appendix II for original references).

stage have dissipated, the projectile and a comparably sized volume of the target are immersed in a broad zone of nearly uniform pressure. The subsequent cratering flow depends more on the pressure and size of this zone than on the extreme pressure highs of jetting.

Numerical computations have shown that the mean pressure during the contact and compression stage can be accurately estimated by a technique called the *planar impact approximation*, sometimes also called *impedance matching*. Most of the projectile and comparable masses of the target are shocked to this pressure. The mean pressure sets a practical maximum for pressures during the subsequent shock wave expansion and is one of the principal controls on how much of

the target is either melted or vaporized in the succeeding excavation stage. The planar impact approximation also permits estimates of energy partition during the earliest stages of impact and provides some insight into projectile penetration and the development of the detached shock structure.

#### 4.5.1 The planar impact approximation

Figure 4.5 illustrates the principal features of the planar impact approximation. The projectile is represented by an infinite sheet of projectile material with thickness equal to the projectile diameter  $L$ . The target is represented by a half space, again composed of the same material as the actual

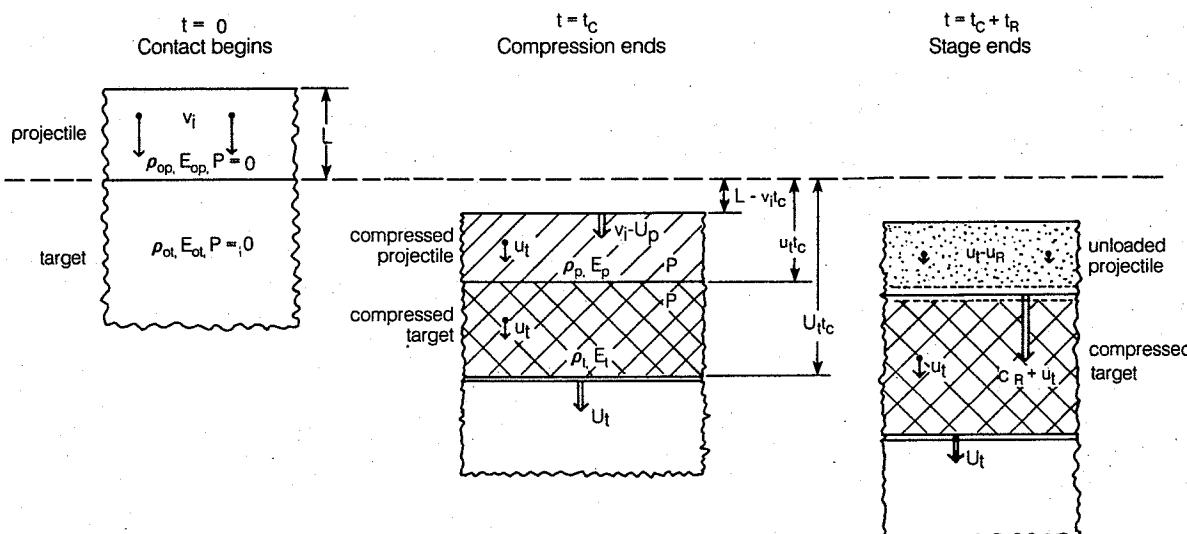


Fig. 4.5 The planar impact approximation in three phases of the contact and compression stage. The projectile, represented by an infinite slab of thickness  $L$  moving at speed  $v_i$ , just contacts the target's surface at  $t = 0$ . Strong shocks propagate into both target and projectile, bringing them to a common particle velocity  $u$ , and pressure  $P$ . The compression phase ends at  $t = t_c$ , when the shock in the projectile just reaches its rear surface. The projectile-target interface at this time is a distance  $u_i t_c$  below the original ground level. A rarefaction, starting from the rear of the projectile, just reaches the projectile-target interface at  $t = t_c + t_R$ , ending the contact and compression stage. The unloaded projectile either continues downward at reduced speed or, especially when vaporization occurs, expands in an upward direction.

target. The sheet, which extends laterally to infinity, slaps vertically against the target at the actual impact velocity  $v_i$ . The impact problem, two- or three-dimensional in reality, is thus reduced to a one-dimensional approximation that can be readily solved. This approximation neglects the rarefactions that propagate inward from the sides of the actual projectile and so tends to overestimate the actual mean pressure. Rarefactions become important when the shock wave in the target has propagated about one projectile diameter. The approximation is thus adequate to assess the initial coupling of energy from the projectile to the target and examine some of the early-time high pressure effects. Subsequent events (the excavation stage) must be analyzed by different techniques.

Immediately after first contact, two shocks propagate away from the interface, one traveling into the projectile and the other traveling downward into the target. Between the shocks the target and projectile are both raised to the same high pressure and must travel together with the same particle velocity, since neither interpenetration nor separation can occur. Other properties, such as the compressed densities of the projectile and target material, their internal specific energies, and the velocities of the shock waves may differ for the two materials.

The conditions between the shocks are determined by applying the Hugoniot equations (3.4.1 to 3.4.3) to each shock separately, using the appropriate equation of state (Eq. 3.4.4) for each material and enforcing the equality of pressure and particle velocity at the interface. Care must be taken to apply the Hugoniot equations to the projectile in a coordinate frame in which it is at rest, since the equations as presented here implicitly assume that unshocked material is at rest. The unshocked projectile is actually moving at  $v_i$  in the target's rest frame, so that if  $u_p$  is the particle velocity behind the shock in the projectile's rest frame, then the same material moves at speed  $(v_i - u_p)$  from the target's point of view. Similarly, if the shock wave in the projectile moves at speed  $U_p$  with respect to unshocked projectile material, then it moves at  $(v_i - U_p)$  with respect to the target (see Fig. 4.5). The equality of particle velocity in both projectile and target thus requires

$$u_t = v_i - u_p \quad (4.5.1)$$

where  $u_t$  is the change in particle velocity across the shock in the target.

Equation 4.5.1 is the basis of a useful graphical method to find the pressure between the shocks in Figure 4.5. The Hugoniot curve, normally plotted on the  $P$ - $V$  diagram or, equivalently, on a  $U$ - $u_p$  plot, may also be drawn on a plot with axes  $P$ , pressure, and  $u_p$ , particle velocity. These three kinds of plot may be freely interconverted by Equations 3.4.5 to 3.4.8. The utility of this plot is shown in Figure 4.6 where the Hugoniot curve for

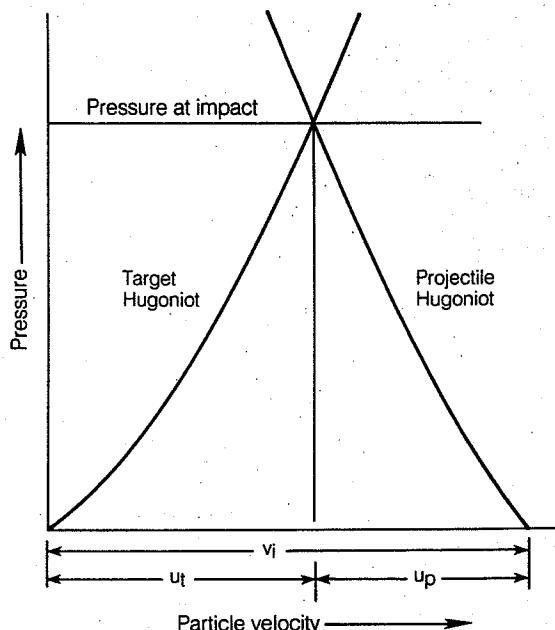


Fig. 4.6 Graphical technique for computing the shock pressure in the planar impact approximation. The target Hugoniot curve is plotted normally as a function of pressure and particle velocity. The projectile Hugoniot is plotted backwards, starting at a particle velocity equal to the impact velocity  $v_i$ . The curves' intersection gives the pressure behind the shock waves in both the target and projectile. The particle velocity in the target  $u_t$  and projectile  $u_p$  can also be read off from the intersection point.

the target is plotted normally as a function of  $u_p$ . The Hugoniot curve for the projectile, however, is plotted backwards, starting from zero pressure at velocity  $v_i$ . The pressures in both the target and projectile are equal where the two curves cross. The particle velocities in both the target and the projectile can be read directly off the graph. The graph's construction guarantees that Equation 4.5.1 is satisfied. Once the pressure or a particle velocity, say  $u_p$ , is known at the intersection, all other quantities of interest can be computed using the equations of state and the Hugoniot equations.

This graphical method, or techniques that do numerically what Figure 4.6 illustrates graphically, can be used with arbitrarily complicated Hugoniot curves. It is easily used with the Tillotson equation since the function  $P(u_p)$  is readily computed for both the target and the projectile (see Appendix II for a discussion of this equation of state and for parameters describing a number of geologic materials). It is seldom possible to obtain an algebraic solution for the pressure. There is one important case, however, where such a solution exists.

It was noted in Section 3.4 that many materials

## IMPACT CRATERING

show a linear relation between shock velocity  $U$  and particle velocity  $u_p$  (Eq. 3.4.9) that often holds to surprisingly high pressures. If both the target and projectile can be described by a linear relation of this form, an algebraic solution may be obtained for the pressure, particle velocities, and other quantities of interest in the planar impact approximation. The relations between shock velocities in the target and projectile,  $U_t$  and  $U_{p_t}$ , respectively, and the corresponding particle velocities,  $u_t$  and  $u_{p_t}$ , are thus assumed to be

$$U_t = C_t + S_t u_t \quad (4.5.2)$$

$$U_{p_t} = C_{p_t} + S_{p_t} u_{p_t} \quad (4.5.3)$$

Where  $C_t$ ,  $C_{p_t}$  (dimensions of velocity) and  $S_t$ ,  $S_{p_t}$  (dimensionless) are empirically determined parameters that describe the target and projectile materials. The uncompressed densities of the target and projectile are  $\rho_0$  and  $\rho_{0p}$ , respectively.

Once one particle velocity, say  $u_t$ , is known, then the other can be computed from Equation 4.5.1. The shock velocities in both materials are, in turn, computed from Equations 4.5.2 and 4.5.3. The shock pressure  $P$ , compressed densities  $\rho_t$  and  $\rho_{p_t}$ , and specific internal energies  $E_t$  and  $E_{p_t}$  are all computed from the Hugoniot equations (3.4.1 to 3.4.3). The particle velocity  $u_t$  is a solution of the quadratic equation that results from

substituting Equations 4.5.2 and 4.5.3 into 3.4.1, using 4.5.1.

$$u_t = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \quad (4.5.4)$$

The constants  $A$ ,  $B$ , and  $C$  are defined in terms of the impact parameters:

$$A = \rho_0 S_t - \rho_{0p} S_{p_t} \quad (4.5.5)$$

$$B = \rho_0 C_t + \rho_{0p} C_{p_t} + 2\rho_{0p} S_{p_t} v_i \quad (4.5.6)$$

$$C = -\rho_{0p} v_i (C_{p_t} + S_{p_t} v_i) \quad (4.5.7)$$

This simple formula gives a good approximation to the conditions in the first stages on an impact. Values of the material constants  $\rho_0$ ,  $C$ , and  $S$  for a number of geologic materials are listed in Appendix II. Table 4.2 uses the planar impact approximation and some of the equations derived later in this section to estimate a number of the parameters that characterize the contact and compression stage.

Note that in the event that the projectile and target are composed of the same material, the denominator  $A$  is zero, apparently leading to trouble with Equation 4.5.4. In reality, the numerator also vanishes when  $A = 0$ , but the ratio is finite and approaches  $-C/B$ , which equals  $\frac{1}{2} v_i$  in this limit. The reader must provide a special exception for

TABLE 4.2 Planar impact approximation results  
Iron projectile\* on Gabbroic anorthosite† (hpp) target

Parameter	Impact Velocity (km/second)		
	5	15	30
Pressure, GPa‡	120.	602.	1920.
<i>Projectile Parameters:</i>			
Particle velocity, $u_p$ km/second	2.17	6.10	11.9
Shock velocity, $U_p$ km/second	7.10	12.7	20.8
Density, $\rho$ Mg/m <sup>3</sup>	11.2	15.0	18.2
Specific energy, E J/kg	2.35	18.6	70.8
Phase on release	solid	part. vapor	vapor
Rarefaction speed, $c_R$ km/ second	7.8	14.0	22.3
Energy partition, $f_p$ (KE)	0.32	0.35	0.36
$f_p$ (IE)	0.19	0.17	0.16
<i>Target Parameters:</i>			
Particle velocity, $u_t$ km/second	2.83	8.90	18.1
Shock velocity, $U_t$ km/second	10.7	17.1	26.8
Density, $\rho$ Mg/m <sup>3</sup>	5.39	8.27	12.2
Specific energy, E J/kg	4.00	39.6	164.0
Phase on release	inc. vapor	vapor	vapor
Rarefaction speed, $c_R$ km/second	10.7	16.2	22.9
Energy partition, $f_t$ (KE) = $f_t$ (IE)	0.25	0.24	0.24

\*Iron parameters:  $C = 4.05$  km/second,  $S = 1.41$ ,  $\rho_0 = 7.8$  Mg/m<sup>3</sup>.

†Gabbroic anorthosite (hpp) parameters:  $C = 7.71$  km/second,  $S = 1.05$ ,  $\rho_0 = 3.965$  Mg/m<sup>3</sup>.

‡These pressures differ from Table 4.1 because the lpp-hpp transition is not included here, whereas it is included in Table 4.1.

this case if Equations 4.5.4 to 4.5.7 are implemented in a computer program.

#### 4.5.2 Energy partition during compression

The projectile's initial kinetic energy is divided or partitioned between several reservoirs during compression. This energy distributes itself over both the projectile and target. It either becomes internal energy, which can do further work during expansion and initiate phase changes, or remains kinetic.

The planar impact approximation permits estimates of the initial energy partition. Subsequent partition during excavation is treated in Section 5.2. Figure 4.5 illustrates the geometrical relations among the various energy reservoirs.

Compression ends when the shock in the projectile just reaches its rear surface, at time  $t_c = L/U_p$ . At this time the shock in the target has reached a depth of  $U_t t_c$ , the projectile-target interface has reached  $u_t t_c$ , and the rear of the projectile is at a depth of  $L - u_t t_c$  which may be either positive (below the original surface) or negative (above it), depending upon the projectile equation of state. Energy partition in the planar impact approximation is performed on a per-unit-area basis. Thus, the projectile's initial kinetic energy is  $\frac{1}{2} \rho_{0p} L v^2$ , per unit area. The target's internal and kinetic energy are both equal to  $\frac{1}{2} \rho_{0t} U_t t_c u^2$ . The Hugoniot equations can be used to show that the fraction of the projectile's initial kinetic energy that is partitioned into the target as kinetic energy,  $f_p(KE)$ , or internal energy,  $f_p(IE)$ , is

$$f_p(KE) = f_p(IE) = (u_p/u_i)/(1 + u_p/u_i)^2 \quad (4.5.8)$$

All dependence upon unit areas is divided out when the energy partition is thus expressed as a fraction.

The fraction of kinetic energy remaining in the projectile at the end of compression,  $f_p(KE)$ , and its internal energy  $f_p(IE)$ , which are not equal to one another, are similarly

$$f_p(KE) = 1/(1 + u_p/u_i)^2 \quad (4.5.9)$$

$$f_p(IE) = (u_p/u_i)^2/(1 + u_p/u_i)^2 \quad (4.5.10)$$

When both projectile and target are composed of the same material,  $u_i = u_p$ , and all four fractions are equal to 0.25. Half of the total energy is thus partitioned into the target and half remains in the projectile. Even when the projectile and target are iron and basalt, respectively, energy is partitioned nearly equally into each of the four categories. This observation led Gault and Heitowitz (Sec. 5.1.5) to propose a model of shock wave attenuation that implicitly assumes a 50/50 initial partition into projectile and target, with the target's portion being exclusively responsible for excavating the crater. More recent numerical computations, however, show that projectile energy

may also participate in crater excavation, so the significance of this early partition is less clear.

Figure 5.5, in the next chapter, illustrates energy partition as a function of time in the vertical impact of an anorthosite projectile on a similar target at 30 km/second. The initial kinetic energy of the projectile is divided among the various reservoirs within a few units of  $\tau$  (defined in Equation 4.2.1), that is, during the contact and compression stage. Although energy exchange among the reservoirs continues through the excavation stage, the planar impact approximation prediction that 25 percent of the energy goes to each category is roughly verified at times of 2 to  $3\tau$ . The largest deviation, that only 15 percent of the energy is transferred to internal energy in the projectile, is mainly due to rarefactions propagating laterally into the projectile—a factor neglected by the planar impact approximation. During the subsequent excavation stage, more of the projectile's energy is coupled into the target, until the projectile is ultimately left with only about 10 percent of its initial energy, equally divided between kinetic and internal energy.

#### 4.5.3 Unloading of the projectile

The shock wave traversing the projectile reaches its rear surface at the end of compression and is reflected back into the compressed projectile material as a rarefaction or release wave. Spallation, discussed in Section 5.3, may eject a small fraction of the projectile's rear as lightly shocked, finely comminuted debris, but most of the projectile unloads from high pressure and, at common solar system impact velocities, expands into the vapor phase. The projectile in Figure 4.1d is thus expanding rapidly and will eventually leave the crater as an iron vapor plume.

The rarefaction wave's speed is generally so high that it overtakes the shock in the target and rapidly weakens it. Although a full equation of state is necessary to compute the rarefaction speed precisely, it can be estimated from a simplified Murnaghan-type equation of state (Appendix II), using Equation 3.4.11 to derive the sound velocity  $c_R$ . This is the velocity of the head of the rarefaction: the propagation speed depends upon the pressure and falls to common sonic velocities in the low pressure tail. Thus, the rarefaction wave is approximately

$$c_R = [(K_0 + nP)/\rho]^{1/2} \quad (4.5.11)$$

where the projectile bulk modulus  $K_0 = \rho_{0p} C_p^2$ , the dimensionless constant  $n = 4S_p - 1$ , and  $\rho$  is the compressed density of the projectile. Equation 4.5.11 also gives the rarefaction speed in the target if the constants  $K_0$  and  $n$  are computed using the target's parameters  $C_p$ ,  $S_p$ , and  $\rho_{0t}$ . Note that  $c_R$  is the rarefaction speed in the rest frame of the compressed material. Since this material is moving

downward at  $u$ , with respect to the target, the rarefaction moves downward at the sum of the two velocities,  $c_R + u$ , in the target's rest frame.

The unloading phase begins when the shock wave arrives at the rear of the projectile and ends when the rarefaction reaches the projectile-target interface. The rarefaction, traveling at speed  $c_R$  with respect to the projectile, must traverse its compressed thickness  $(\rho_0/\rho)L$  and so the duration of the unloading phase  $t_R$  is

$$t_R = (\rho_0/\rho)L/c_R \quad (4.5.12)$$

The unloaded projectile material accelerates upward during its release. If the maximum shock pressure is low and the velocity doubling rule is approximately valid, the unloaded projectile's particle velocity with respect to the target is  $u - u_p$ . This is nearly zero for a projectile and target of identical composition, but can be either upward or downward if their Hugoniot curves differ. When shock pressures are high enough to vaporize the projectile, the vapor may acquire a net *upward* velocity component, expanding out of the crater as a hot, high-speed vapor plume. Its velocity is determined by Equation 3.4.12, which requires detailed knowledge of the projectile equation of state. However, if the internal energy  $E$  is much larger than the energy of vaporization  $E_{vap}$ , then the difference is available for conversion to kinetic energy of the gas. The ultimate root-mean-square vapor plume velocity is  $u_R = \sqrt{2(E - E_{vap})}$  with respect to the projectile, or  $u_R - u$  (upward) with respect to the target. The mechanics of vapor plume expansion is considered in more detail in Section 5.3.

#### 4.5.4 Limitations of the planar impact approximation

The planar impact approximation gives good estimates of the mean shock pressures that develop during the short-lived contact and compression stage, the early energy partition between the target and projectile, and the initial particle velocities, shock speeds, and rarefaction velocities. These estimates hold best along the central axis of a vertical impact.

Even in a vertical impact, however, the sides of the projectile are affected by rarefactions propagating inward from free surfaces closer than the rear of the projectile. These rarefactions modify the course of compression, leading to lower mean pressures and smaller downward velocities. At the same time, the presence of lateral pressure gradients produces radially outward velocity components. The center of the projectile thus moves downward into the target at speeds given by the planar impact approximation, while its sides move radially outward as well as downward at reduced velocity. The resulting velocity field "turns the projectile inside out," as in Figures 4.1c and

4.1d, and the projectile is smeared out along the inside of the rapidly expanding crater cavity.

The planar impact approximation's applicability to highly oblique impacts is unknown. Perhaps the approximation describes the average pressures if the vertical velocity component is used instead of the impact velocity itself. However, jetting may be inseparable from the early stages of compression and no simple approximate scheme exists. In any event, more work on oblique impacts is needed.

#### 4.6 THE END OF CONTACT AND COMPRESSION

The contact and compression stage is considered to end when the rarefaction unloads the projectile. This choice is somewhat arbitrary: some authors consider that it ends when the shock wave reaches the rear of the projectile. However, the characteristic feature of the following excavation stage is the expansion of a roughly hemispherical shock wave away from a nearly uniform pressure core region. Since the planar impact approximation seems to work reasonably well through the projectile unloading, and since the shock attains its detached structure only somewhat after compression is complete, projectile unloading is here considered part of contact and compression. Figure 4.5 shows that the stage's end coincides with the development of the detached shock (a spreading, isolated zone of high pressure) in the target.

Another reason for including projectile unloading in the contact and compression stage is that while the projectile's rear and center unloads, its front maintains the projectile-target interface at high pressure, continuing to transfer momentum and energy into the target. This transfer ceases only when the rarefaction arrives at the interface. After this, the projectile more or less "goes along for the ride" on the inside of the expanding crater cavity or even expands out of it as a vapor plume.

The total duration of the contact and compression stage is given precisely in the planar impact approximation as the sum of the durations of compression and unloading,  $(t_c + t_R) = L/U_p + (\rho_0/\rho)L/c_R$ . The actual duration is somewhat longer in numerical computations that take into account the geometry and possible obliquity of the impact. An estimate of a few times  $\tau$  from Equation 4.3.1 is probably realistic, as well as straightforward to calculate.

The high pressures of the contact and compression stage decline rapidly in the subsequent excavation stage. The shock wave expands rapidly, engulfing, compressing, and accelerating more target material as it weakens. After the shock wave dissipates, the residual velocity field opens the crater. The planar impact approximation is not valid during the excavation stage, and material strength

and the target's gravitational field cannot be neglected. Although much attention was focused on the projectile during contact and compression, the target becomes the center of activity during excavation. New sets of approximations are needed to describe the cratering process during this stage, marking a natural separation from contact and compression.

### FURTHER READING

#### *General References*

The best general, although brief, exposition of the mechanics of impact cratering is

D. E. Gault, W. L. Quaide, and V. R. Oberbeck, Impact cratering mechanics and structures, in B. M. French and N. M. Short (Eds.), *Shock Metamorphism of Natural Materials*, Mono Book Co., Baltimore, Md., pp. 87-99 (1968).

Although a number of other good reviews of impact cratering mechanics exist, they are in hard-to-find limited circulation reports. The reader can find a detailed discussion of the contact and compression stage in Kieffer and Simonds (1980) cited at the end of Chapter 3. An older, but historically very important, discussion of impact cratering mechanics is

E. M. Shoemaker, Impact mechanics at Meteor Crater, Arizona, in B. M. Middlehurst and G. P. Kuiper (Eds.), *The Solar System*, 4: 301-336 (1963).

The theory of explosion cratering is closely related to that of impact cratering. A theory of explosion cratering that attributes much of the early-time motion to spalls ejected from the target surface is given by

M. D. Nordyke, Nuclear craters and a preliminary theory of the mechanics of explosive crater formation, *J. Geophys. Res.* 66: 3439-3459 (1961).

#### *Special Topics*

There is very little experimental information on the contact compression stage, largely because of its short duration with laboratory-size projectiles. There are a number of numerical computations that illustrate this stage, mostly for vertical impacts. The best and most readily available of these is

J. D. O'Keefe and T. J. Ahrens, Shock effects from a large impact on the moon, *Proc. Lunar Sci. Conf.* 6<sup>th</sup>, pp. 2831-2844 (1975).

The computations described in the above paper were later revised when a high-pressure phase transition was discovered in the gabbroic anorthosite target material. The importance of such phase transitions is well brought out in the subsequent paper:

J. D. O'Keefe and T. J. Ahrens, Impact flows

and crater scaling on the moon, *Phys. Earth and Planet. Int.* 16: 341-351 (1978).

There is almost nothing yet in print on the contact and compression stage for oblique impacts. Jetting, on the other hand, has a fairly large literature. The most useful papers from the viewpoint of impact cratering include the classic description of the hydrodynamic theory:

G. Birkhoff, D. P. MacDougall, E. M. Pugh, and G. Taylor, Explosives with lined cavities, *J. Appl. Phys.* 19: 563-582 (1948).

The most complete summary of the theory of jetting, including a discussion of the limiting angle and asymmetric jetting is

J. M. Walsh, R. G. Shreffler, and F. J. Willig, Limiting conditions for jet formation in high velocity collisions, *J. Appl. Phys.* 24: 349-359 (1953).

A particularly clear numerical study of symmetric jetting is

F. H. Harlow and W. E. Pracht, Formation and penetration of high-speed collapse jets, *Phys. Fluids*, 9: 1951-1959 (1966).

Conditions for jetting are derived for silicates by S. W. Kieffer, Impact conditions required for formation of melt by jetting in silicates, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 751-769 (1977).

The same author also gives a clear, nonmathematical summary of jetting, discusses jetting as a source of internal melt pockets in shocked porous rocks, and proposes a jetting origin for melts that cool to form one class of chondrules in

S. W. Kieffer, Droplet chondrules, *Science* 189: 333-340 (1975).

The classic paper on the planar impact approximation was published only in an obscure symposium proceedings, but has been so influential that it is widely cited. Copies of the proceedings can often be found in Aeronautical Engineering library collections:

D. E. Gault and E. D. Heitowitz, The partition of energy for hypervelocity impact craters formed in rock, *Proc. Sixth Hypervelocity Impact Symp.* 2: 419-456 (1963).

This important paper is cited in the next chapter in connection with shock-wave attenuation. This paper also contains the first report of jetting in impacts, a process described in more detail by Gault et al. (1968), cited earlier.

The planar impact approximation was used with more complex equations of state by

E. M. Shoemaker, Penetration mechanics of high velocity meteorites, illustrated by Meteor Crater, Arizona, in *Rept. of the Int. Geol. Congress*, XXI Session, Norden, Copenhagen, Part XVIII, pp. 418-434 (1960).

## V. CRATERING MECHANICS: EXCAVATION STAGE

### 5.1 INTRODUCTION

The excavation stage is characterized by two major processes that coexist throughout the stage at different degrees of completion. The earliest process is the expansion of the shock wave produced during the previous contact and compression stage. The shock wave expands and weakens, degrading to a stress wave as it engulfs progressively more target material. It decays in strength both because its energy is spread over a greater volume of material and because of irreversible energy losses in the shock-compressed material. The stress wave travels quickly, never moving slower than the bulk sound speed in the target. As it passes, the wave sets target material in motion, paving the way for the second major process, the excavation flow. The excavation flow actually opens the crater. This subsonic flow of the target material occurs on a much longer timescale than stress wave expansion: the stress wave is usually long gone, decayed into elastic waves, by the time the excavation flow has completely formed the transient crater.

These two processes are responsible for different physical phenomena. Thus, the expansion and decay of the shock wave is the main regulator of thermodynamic processes during cratering. The masses and temperatures of melted and vaporized material are determined almost exclusively by the dynamics of the expanding shock wave. On the other hand, the shape and size of the final crater are almost exclusively determined by the details of the excavation flow and its interaction with the target planet's strength and gravity. Only a few aspects of cratering, such as vapor plume expansion and near-surface spallation, involve aspects of both stress wave expansion and excavation flow.

This natural division of major processes suggests that the excavation stage might be divided into parts emphasizing each process. However, the two processes are so intimately related that such a division would be deceptive: the stress wave actually initiates the excavation flow. Furthermore, the excavation flow near the impact site begins before the stress wave has set more distant material into motion. The major activity times of these processes thus overlap in a cratering event. Although the stress wave and excavation flow are spatially distinct, the boundary between them moves according to a complex law, making a clean separation impossible.

Shock wave attenuation and the excavation flow must thus be treated in a unified way, as part of the same stage of cratering. The following sections, however, discuss their distinctive features separately, emphasizing approximations and modes of description best suited to each process. The links between the two processes are emphasized where appropriate. Vapor plume expansion and spallation, phenomena that partake somewhat of each major process, are discussed in their own small sections. It is hoped that this arrangement will give the reader a clear perception of the complex and interrelated events occurring in the excavation stage.

### 5.2 SHOCK WAVE EXPANSION AND ATTENUATION

#### 5.2.1 Shock (stress) wave structure

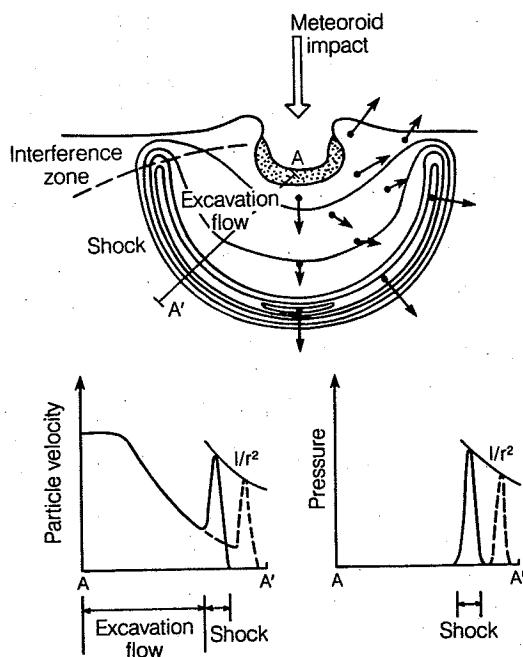
Figure 4.1d in the last chapter illustrates the structure of the shock wave at the beginning of the excavation stage for a vertical impact. Much less is known about oblique impacts. It appears that the shock wave in the target is broadly similar, but the shock is weaker and more diffuse than that of a vertical impact at the same velocity. The predominant feature of the shock wave throughout the excavation stage is that it is *isolated*, or *detached*. This detachment, discovered during numerical code calculations in the mid-1960s (the first published mention I know of was in Bjork et al., 1967), begins during the contact and compression stage when the rarefaction from the rear of the projectile encroaches on the downward-moving shock wave in the target.

The detached shock expands away from the impact site, maintaining an approximately constant thickness as it weakens. Figure 5.1 illustrates the detached shock configuration a short time after the impact. The contours outline a nearly hemispherical shell of high pressure, centered on a point roughly one projectile diameter below the initial ground surface. Shock pressures are largest directly below the impact site, although they do not vary greatly over the hemisphere. The exception is in the near-surface interference zone where the pressure drops rapidly with decreasing depth, reaching zero at the surface itself. The insets accompanying Figure 5.1 show the pressure and particle velocity through the shock along the sec-

Fig.  
sho  
teor  
the  
time  
pres  
an  
atot  
inse  
tim  
the  
of t  
pea

tior  
are  
Hu  
nea  
loci  
thir  
resi  
ver:  
and

I  
hov  
hav  
sho  
bot  
cre  
late  
lop  
in 1  
larg  
and  
wh  
the



**Fig. 5.1** Schematic illustration of the expanding shock wave and excavation flow following a meteorite impact. The contours in the upper part of the figure represent pressure at some particular time after the impact. The region of high shock pressure is seen to be isolated or "detached" on an expanding hemispherical shell. The insets show profiles of particle velocity and pressure along the section AA'. The dashed lines on these insets show particle velocity and pressure some time later than those shown by the solid lines, and the solid curves connecting the peaks are portions of the "envelopes" of peak particle velocity and peak pressure.

tion line AA'. Both pressure and particle velocity are high in the shock, where they are given by the Hugoniot equation (3.4.2). The pressure drops to nearly zero behind the shock, but the particle velocity declines to a residual level of about one-third to one-fifth of its peak at a given radius. This residual velocity is due to thermodynamic irreversibility in the shock, described in Section 3.4.3, and marks the beginning of the excavation flow.

The dashed lines in insets in Figure 5.1 show how the pressure and particle velocity profiles have changed at a time somewhat later than those shown by solid lines. As the shock wave expands, both peak pressure and peak particle velocity decrease roughly as  $1/r^2$  (although, as will be seen later, pressure may fall more steeply). The envelopes defined by these peak values, also sketched in the insets, are important because they give the largest value ever attained at a given radius. A number of figures presented later in this chapter are envelopes of this kind. One should use care when interpreting such curves, however, because the peak values are not achieved simultaneously

at all radii. The structure of the shock wave cannot be inferred from the envelopes of maximum pressures or particle velocities.

The shock front is not a mathematical discontinuity as its name might imply: the pressure rises from near zero in front of the shock to its peak value over a finite rise time  $\tau_r$ . The pressure likewise declines behind the shock over a finite interval, the decay time, here denoted  $\beta \tau_r$ , where the factor  $\beta$  is usually larger than one. There are two reasons why the shock wave produced by an impact is not discontinuous. First, rate-dependent processes in the target material broaden the shock wave. Although the shock thickness may be only millimeters in metals, the wide variety of failure processes occurring in rocks (such as pore collapse, yielding at the Hugoniot Elastic Limit, phase transformations, and scattering) produce shock wave thicknesses of tens of meters, with corresponding rise times of milliseconds. Second, an irregular projectile takes a certain amount of time to fully contact the target, long enough that nonlinear processes are unable to greatly steepen the shock front in a spherically expanding wave. On this basis alone the rise time of the shock or stress pulse generated by the vertical impact of a spherical projectile of diameter  $L$  at velocity  $v_i$  is

$$\tau_r = L/2v_i \quad (5.2.1)$$

For projectiles large enough to be of geological interest the rise times given by Equation 5.2.1 are much longer than those given by rate-dependent processes, which are therefore not included in subsequent discussions.

The rise times of the stress waves produced by impacts have not received much attention, either experimentally or in numerical computations, because it was believed that rise time plays only a minor role in the cratering process. Although this is true so far as maximum pressures and excavation flow velocities are concerned, the rise time plays a critical role in defining the size of the near-surface zone and the maximum stresses within it, as well as the sizes of fragments produced by the stress wave. Lacking definite information to the contrary, Equation 5.2.1 will be used to estimate rise times where they are needed. Although this choice is consistent with the results of the small number of code computations that do not deliberately smear the stress pulse, more research is needed on stress wave rise times in impacts.

The shock wave continues to expand throughout the excavation stage, weakening with time. It first loses its status as a supersonic shock, degrading into a strong stress wave with an elastic precursor. It eventually decays into an elastic wave that carries only about  $10^{-4}$  of the original impact energy away from the site. The rate of decline of the shock wave's strength determines the mass of melted or vaporized target material. The factors

that govern shock wave attenuation are examined in detail in the next section.

### 5.2.2 Pressure and particle velocity attenuation

Figure 5.2 illustrates the detached shock or stress wave schematically and defines parameters that describe it. The shock wave attains this structure before it has propagated more than one or two projectile diameters from the impact site, as seen in Figure 4.1d. The shock or stress wave (the elastic precursor, which is energetically insignificant, is ignored throughout this discussion) can be roughly characterized by a mean pressure  $\langle P \rangle$  and particle velocity  $\langle u_t \rangle$  that are constant across the effective width  $w_s$  of a hemispherical shell of outer radius  $r$ . The shell travels with the shock (or plastic) velocity  $U_t$ .

The radial momentum  $p_s$  carried by the detached shock, neglecting the near-surface zone, is approximately given by

$$p_s = 2\pi\rho[r^2 - (r - w_s)^2]\langle u_t \rangle \quad (5.2.2)$$

Similarly, the total energy  $E_s$ , kinetic plus internal, in the detached shock is

$$E_s = 2\pi\rho[r^2 - (r - w_s)^2]\langle u_t^2 \rangle \quad (5.2.3)$$

It is important for the ensuing discussion to note that the mean particle velocity in the shock,  $\langle u_t \rangle$ , and the square root of the mean square particle velocity,  $\sqrt{\langle u_t^2 \rangle}$ , may be very different, especially if the particle velocity reverses sign in the wave, as it does in elastic waves.

If the shock width  $w_s$  is nearly constant throughout its expansion, and if the compressed density  $\rho$  is not too different from the uncompressed density, then it is readily seen that the particle velocity in a wave that conserves radial momentum falls roughly as  $1/r^2$ . The mean pressure  $\langle P \rangle = \rho_0 U_t \langle u_t \rangle$  in this wave falls more steeply than  $1/r^2$  because  $U_t$  decreases as the pressure falls. The particle velocity in a hypothetical wave that conserves energy alone falls roughly as  $1/r$ . The momentum in such a wave, far from

being conserved, actually *increases* as the radius of the wave increases.

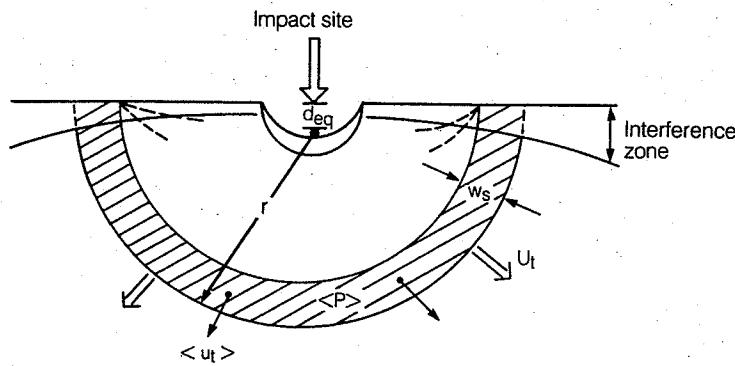
Elastic waves, which conserve both energy and momentum, satisfy Equations 5.2.2 and 5.2.3 simultaneously because they develop a negative particle velocity tail where material springs back toward the center of expansion. This rebound causes the mean particle velocity  $\langle u_t \rangle$  to fall as  $1/r^2$  while the root mean square particle velocity  $\sqrt{\langle u_t^2 \rangle}$  falls only as  $1/r$ . This negative velocity tail develops when the previously compressed material expands, overshooting the decompressed state in an "elastic rebound." It occurs only when the associated tensile stress,  $\sigma_{el} = \rho_0 c_L u_t$ , is smaller than the sum of the target's tensile strength plus any overburden stress that may be present. If the wave is much stronger than this limit, the rebound induces tensile rock fracture and the wave becomes inelastic.

Since there are few ways to lose momentum, other than to an unusually strong excavation flow, a strong inelastic wave tends to conserve its momentum while its energy declines. This tendency of strong waves to conserve momentum is confirmed by the study of stress waves generated by underground nuclear explosions. A data compilation by Perret and Bass (1975) shows that in the "hydrodynamic" regime of particle velocities ranging from 30 m/second up to 5 km/second, the peak particle velocity  $u_{max}$  declines as

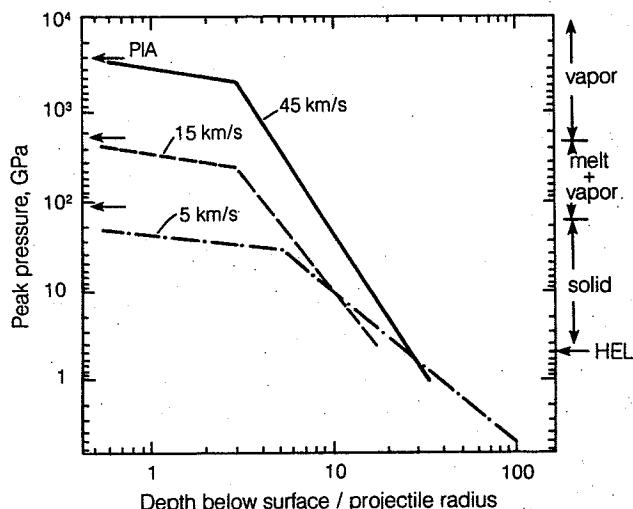
$$u_{max} \propto r^{-1.87 \pm 0.05} \quad (5.2.4)$$

where  $r$  is the distance from the nuclear explosive. The peak particle velocity is the maximum particle velocity recorded at any time at a given range  $r$ . It is thus essentially the same as the particle velocity  $u_t$  in the detached shock. Equation 5.2.4 holds equally well in alluvium, tuff, and granite. Among these diverse materials the deviation of  $u_{max}$  from Equation 5.2.4 is no larger than a factor of 1.7.

The displacements in the target caused by the shock itself are small. The radial displacement  $d_r$  is the integral of the velocity over time,  $d_r =$



**Fig. 5.2** Variables describing the expanding shock wave are the radius  $r$  of the shock front away from the effective center at depth  $d_{eq}$ . The shock of width  $w_s$  moves with velocity  $U_t$  in the target. The mean pressure in the expanding shock wave is  $\langle P \rangle$  and the mean particle velocity is  $\langle u_t \rangle$ .



**Fig. 5.3** The decline in peak shock pressure as a function of depth below the surface for an iron meteorite striking an anorthosite target at 5, 15, and 45 km/second. Note the log-log scales. At first, the pressure falls slowly from the maxima given by the planar impact approximation (arrows marked "PIA"). However, at depths greater than a few projectile radii it falls more rapidly. The rate of fall is faster for the higher velocity impacts. The scale on the right-hand side of the plot shows the Hugoniot Elastic Limit (HEL) of the target and the pressure ranges over which melt and vapor form. *After Ahrens and O'Keefe (1977).*

$\int u dt$ . Integrating only over the shock,  $d_s \approx u_{\max} \tau$ . In the immediate vicinity of the projectile,  $d_s$  is thus of the order of the projectile diameter. Because the shock duration is nearly constant,  $d_s$  falls nearly as  $1/r^2$  at greater distances, following the decline of  $u_{\max}$ , so that these displacements can be neglected more than a few projectile radii from the impact site. Much larger displacements develop during the excavation flow because of its much longer duration. Even though excavation flow velocities are smaller than the particle velocities in the shock, its long duration more than makes up for the low velocities.

Unlike particle velocity, pressure in the detached shock does not follow a simple, nearly universal law of decline. Pressure is related to particle velocity by the second Hugoniot equation (3.4.2), and thence to the equation of state. The pressure in different materials, or in the same material over different pressure ranges, declines at a rate dependent upon the equation of state.

In spite of the simple relation between particle velocity and radius, however, most studies of both impact and explosion cratering concentrate on the pressure decline with radius. The reason for this is probably historical, although E. J. Öpik pointed out the utility of the momentum conservation approach as early as 1958. Öpik, however, did not explicitly derive the  $1/r^2$  particle velocity dependence and his work was ignored. Few, if any, numerical code computations of impacts report peak particle velocity as a function of range from the impact site. A number of explosion computations, however, do show the expected  $1/r^2$  dependence. A fuller appreciation of momentum

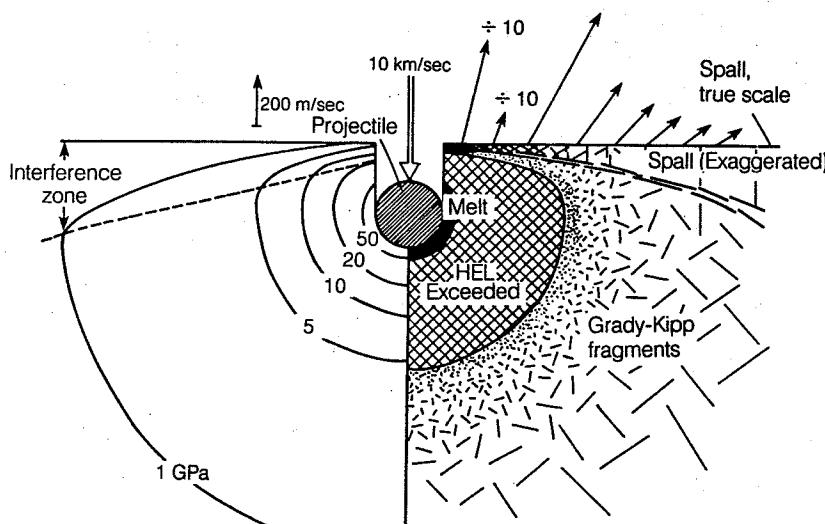
conservation in the detached shock might lead to greater insight into the mechanics of impact cratering.

Figure 5.3 illustrates peak pressure as a function of depth directly beneath the impact site of an iron meteorite that strikes a gabbroic anorthosite target at 5, 15, and 45 km/second. These curves are the envelopes of the peak pressure in the detached shock as it sweeps outward from the impact site. Remember that these pressures are not achieved simultaneously at all radii. They are only the maxima that are achieved at *some* time during the shock wave expansion. The peak pressure is nearly constant close to the origin, declining by at most a factor of two from the planar impact approximation pressures (marked "PIA" in the figure) to a depth of 1.5 to 2 projectile diameters. This is the near field region defined during contact and compression. The rate of pressure decline then increases sharply and remains steep out to the largest depths included in the computation. The computed rate of decline ranges between about  $1/r^{1.5}$  for the low velocity impacts to  $1/r^3$  for the highest velocity impacts. In the elastic limit, reached at large radii, the pressure declines as  $1/r$ .

### 5.2.3 Melting and vaporization

Figure 5.3, besides illustrating the decline of pressure with distance away from an impact, also shows the pressures at which various phase transitions occur in the target. Thus in a 45-km/second impact of an iron projectile, the gabbroic anorthosite target vaporizes to a depth of about 2.5

## IMPACT CRATERING



**Fig. 5.4** Schematic illustration of the falloff of pressure near the site of an impact and its implication for the final state of the target. Beneath the target, pressure contours are approximately hemispherical and grade outward from vapor (if present) to melt, then to the limit of crushing at the Hugoniot Elastic Limit. Near the free surface rarefactions reduce the maximum pressure in the interference zone. The tensile rarefaction wave throws off a thin spall plate from the surface at very high speed, indicated by the arrows. Below the spall zone tensile stresses break the rock into Grady-Kipp fragments to great depths below the impact site. Near the surface these fragments are also expelled at high speed. *After Melosh (1984).*

projectile diameters below the original surface and melts down to about 5 diameters. The total mass of material melted or vaporized is computed by integrating the volume of material enclosed within the appropriate maximum pressure contours. Figure 5.4 schematically illustrates the shape of these contours. Note the rapid decline of pressure near the surface, where the pressure contours must approach zero because of the free-surface boundary condition, and the hemispherical shape of the deeper contours.

The total mass of material melted or vaporized in a gabbroic anorthosite target, in units of projectile mass, is given by Equation 7.10.1 as a function of impact velocity. See Section 7.10 for more information. At low impact velocities, the amount of melt or vapor is sensitive to geometric factors because the shape of the vaporized or melted region is complex, but at higher impact velocities the affected regions are nearly hemispherical and the mass of either melt or vapor is proportional to the square of the impact velocity. The mass of melt is roughly ten times larger than the mass of vapor. This general relation is a simple geometrical consequence of the rate of decline of pressure with radius: a steeper dependence on radius ( $1/r^5$ , say) would yield a larger quantity of vapor relative to melt.

The geometry of the shock wave and the decline of its peak pressure are the major factors in defining the volume and distribution of shocked material. The shock wave moves so much faster than the materials left behind it that they may be regarded as nearly stationary. Only after the shock

wave has passed can these shocked materials participate in the excavation flow and mix with debris of different shock levels. This situation suggests that the thermodynamic state of material near the site of an impact can be estimated from some approximate scheme that treats only the expanding shock wave while ignoring the complexities of the excavation flow. So long as the peak pressure beneath the impact site and its rate of decline with radius are known, it is relatively easy to compute the total mass of material shocked to a given pressure and, using the material's thermodynamic equation of state, compute its temperature, how much is in different phases, and many other thermodynamic details. Several such schemes have been derived and are discussed in the next section.

#### 5.2.4 The Gault-Heitowit and other models of shock-wave decay

An ingenious and widely used model of peak shock pressure decay near the site of an impact was devised in 1963 by D. E. Gault and E. D. Heitowit. Their work was based on a still earlier "ballistic pendulum" model for the depth of crater excavation proposed by A. C. Charters and J. L. Summers in 1959. The Gault-Heitowit model begins with an estimate of the initial shock pressure using the planar impact approximation described in Section 4.5.1. Gault and Heitowit then assume that exactly half of the projectile's initial kinetic energy is transferred to the target. (Through an oversight, this fact was not stated explicitly in the

original paper, but its truth is clear from their results and from later public statements by Gault.)

In the simplest version of the Gault-Heitowit model, the energy in the target is assumed to be evenly distributed throughout the hemisphere of radius  $r$  enclosed by the expanding shock wave. The total mass  $M_r$  within this region is

$$M_r = (\frac{4}{3})\pi\rho_{0r}r^3 \quad (5.2.5)$$

where  $\rho_{0r}$  is the density of the unshocked target material (even though the target material is compressed behind the shock, the total mass of material engulfed is given correctly by Equation 5.2.5). The total energy behind the shock at any given time is the specific internal energy  $E_r$ , plus the specific kinetic energy, times the mass enclosed  $M_r$ . Using the Hugoniot equation (3.4.3) for the specific internal energy, which equals the specific kinetic energy, the total energy in the target  $W = \frac{1}{2}mv_i^2$  (where  $m$  is the projectile mass and  $v_i$  is its velocity) is given by

$$W = M_r P(V - V_0) = (\frac{4}{3})\pi r^3 P(1 - \rho_{0r}/\rho_i) \quad (5.2.6)$$

For the special case in which the Hugoniot is a linear function of the compression,

$$P = K(1 - \rho_{0r}/\rho_i) \quad (5.2.7)$$

(which is valid for a pressure wave that moves at constant velocity  $c = (K/\rho_0)^{1/2}$ ), the total energy becomes

$$W = (\frac{4}{3})\pi r^3 P^2/K \quad (5.2.8)$$

Since  $W$  is assumed constant throughout the expansion, this equation predicts

$$P \propto r^{-3/2} \quad (5.2.9)$$

A more complicated pressure dependence on radius is obtained if a more realistic relation between pressure and compression is used. For a linear shock velocity-particle velocity Hugoniot (Equation 3.4.9) the peak pressure declines as  $r^{-3}$  for pressures in excess of 0.1 GPa but then declines as  $r^{-3/2}$  when pressures drop below 0.01 GPa. The high pressure dependence of  $P$  on  $r^{-3}$  is a simple consequence of Equation 5.2.6, because at very high pressures  $V$  is small compared to  $V_0$  or, equivalently,  $\rho_{0r}/\rho_i$  is small compared to 1. When this happens, Equation 5.2.6 requires that  $P \propto r^{-3}$ .

Gault and Heitowit refined this simple model by introducing the concept of *waste heat*. They recognized that not all of the energy behind the expanding shock is available to drive further expansion, since some of it is consumed in heating, melting, or vaporizing the material behind the shock front. To account for this loss they assumed that whenever the shock front expands by a small amount  $\Delta r$ , the total energy  $W$  is decremented by the amount of heat that will eventually be irreversibly deposited in the mass  $2\pi\rho_{0r}r^2\Delta r$  engulfed.

The waste heat is thus subtracted from the total energy budget at the outset and plays no further role in driving the shock wave's expansion. Although a number of authors find this procedure objectionable, it at least ensures that the final energy balance is correct.

The concept of waste heat was used by Gault and Heitowit in conjunction with a simple estimate of its magnitude. Assuming that the release adiabat follows the Hugoniot curve (see Sec. 3.4.3) and that the Hugoniot is described by a linear shock velocity-particle velocity relation (Eq. 3.4.9), the irreversible energy loss (waste heat) per kilogram of the target  $\Delta E_r$  is given by

$$\Delta E_r = (\frac{1}{2})u_i^2[1 - 2[\xi - \xi^2 \ln(1 + 1/\xi)]] \quad (5.2.10)$$

where  $\xi$  is a dimensionless function of the particle velocity in the target,  $u_i$ :

$$\xi = C_i/S_i u_i \quad (5.2.11)$$

where  $C_i$  and  $S_i$  are parameters in the linear shock velocity-particle velocity relation.

This waste heat equation is valid only so long as the release adiabat approximates the Hugoniot curve. Although this is valid at low pressures ( $\lesssim 100$  GPa), at high pressures when vaporization becomes important Equation 5.2.10 greatly overestimates the amount of waste heat and will thus lead to an overly rapid shock-wave decay.

The Gault-Heitowit model including waste heat cannot be solved analytically, but a simple numerical procedure permits computation of the pressure decline as a function of radius. The computation is begun by assuming that a hemisphere of radius  $r_0$  is filled with target material that has been shocked to a uniform high pressure, which is given by the planar impact approximation (Sec. 4.5.1). This initial state is a natural result of the contact and compression stage and is also reflected by numerical models such as that shown in Figure 5.3, in which an inner region of nearly constant pressure develops. This initial hemisphere prevents the  $r^{-n}$  pressure dependence from being a problem at small  $r$ —the model only holds for  $r > r_0$ . The energy  $W$  in this hemisphere is equal to half the initial kinetic energy of the projectile. An initial radial particle velocity is computed from the Hugoniot equation (3.4.3) using the specific internal energy in the hemisphere, which is equal to half the total energy  $W$  divided by the mass contained in the hemisphere. (Remember that energy is partitioned equally between specific internal energy and specific kinetic energy  $u_i^2/2$  in the target, by the Hugoniot equation, 3.4.3, and the relation, 3.4.7.)

The computation is continued by allowing the shock wave to expand in small increments of radius  $\Delta r$ . At each increment a new total available energy  $W(r + \Delta r)$  is computed from the previous energy  $W(r)$  by subtracting the total waste heat,

$2\pi\rho_0r^2\Delta E\Delta r$ . A new particle velocity  $u(r + \Delta r)$  is computed from the new specific internal energy, which is half of the new total energy divided by the total mass engulfed,  $M(r + \Delta r) = (\frac{1}{2})\pi\rho_0(r + \Delta r)^3$ . This new particle velocity is used in the next step to compute the waste heat  $\Delta E$ . This cycle is repeated until some designated radius is achieved. Parameters such as pressure, shock velocity, and density may be deduced at each step from the particle velocity, the Hugoniot equations, and the equation of state.

The rate of pressure decay deduced from the Gault-Heitowit model, however, is *not* in good agreement with the results of numerical code computations. The initial rate of pressure decay predicted by this model is typically proportional to  $r^{-4}$ , which is much steeper than that obtained by more exact techniques, which seldom give rates of decay exceeding  $r^{-3}$ . Shock decay rates deduced from terrestrial impact craters also do not seem to exceed  $r^{-3}$  when the cratering flow field is properly reconstructed (see the discussion in Kieffer and Simonds, 1980). The Gault-Heitowit model, with its too-steep falloff of pressure with distance from the impact, thus tends to underestimate the amounts of melt and vapor produced in an impact. Two factors are probably responsible for the failure of this model. The first, already discussed, is that the amount of waste heat is overestimated by Equation 5.2.10 when melting or vaporization occur and the release adiabat departs significantly from the Hugoniot curve. However, this factor is probably not very important for pressures less than about 100 GPa. Second, the Gault-Heitowit model assumes that the shock energy is evenly distributed over a hemisphere enclosing the impact site.

The current understanding of shock wave geometry (Sec. 5.2.1) shows that the high pressures produced by an impact fill a shell of finite thickness  $w_s$ , rather than a hemisphere. An obvious correction to the Gault-Heitowit approach would thus be to localize the zone of high pressures in a hemispherical shell like that illustrated in Figure 5.2 and repeat the same procedure using the waste heat concept described above. Since the volume over which the energy is distributed depends only upon  $r^2$  for a shell instead of  $r^3$  for a hemisphere, the energy density and hence pressure declines at a slower rate. A disadvantage of this approach is that a new variable, the shock thickness  $w_s$ , is introduced. This variable must somehow be determined. The simplest procedure is to assume that it is constant.

The results of such finite shock thickness models are actually in good agreement with numerical computations such as those illustrated in Figure 5.3, although the usefulness of the approximate models is limited by the necessity of choosing an initial radius  $r_0$  and shock thickness  $w_s$ . An investi-

gation by W. B. McKinnon (Chapman and McKinnon, 1986) yields a rate of pressure decline  $P(r) \propto 1/r^n$ , in which the power  $n$  is given by

$$n = 2 + (f/2)(r_0/w_s - 1) \quad (5.2.12)$$

where  $f$  is the ratio between the actual waste heat to that given by Equation 5.2.10. It must be determined from the release adiabats of the target material. Because  $w_s$  is generally close to  $r_0$ , McKinnon's equation leads to rates of pressure attenuation only somewhat steeper than  $r^{-2}$ .

Although other attempts have been made to find simple ways of describing the rate of pressure decline near an impact, none of these are enough of an improvement over the Gault-Heitowit scheme to repeat here. For many purposes, a sufficient approximation to the rate of pressure decay may be had by substituting the empirical particle velocity decay, Equation 5.2.4, into the Hugoniot equation (3.4.2), determining the shock velocity  $U(r)$  from the linear shock velocity-particle velocity relation, yielding

$$P(r) = \rho_0[C_i + S_i u_0(r_0/r)^{1.87}]u_0(r_0/r)^{1.87} \quad (5.2.13)$$

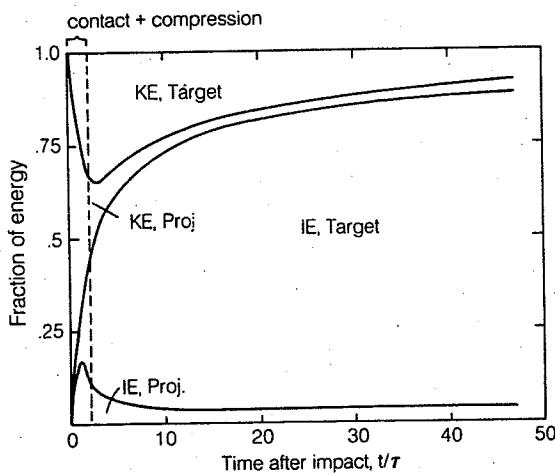
where  $u_0$  is the initial particle velocity at radius  $r_0$ .

More detailed investigation of the pressure decay may require full-scale numerical computations using accurate equations of state. Fortunately, modern computer technology is making this increasingly feasible.

### 5.2.5 Energy partition during crater excavation

The disposition of the projectile's initial kinetic energy as a function of time after the impact is frequently used to describe the course of an impact and to determine how much of this energy eventually ends up as heat buried in the target or in some specific portion of it. Figure 5.5 illustrates the partition of the initial kinetic energy among projectile and target internal and kinetic energy reservoirs. This figure was derived from a numerical computation of a 10-km diameter anorthosite projectile striking an anorthosite target at 30 km/second. The horizontal axis is in units of  $\tau$ , the time it takes the projectile to traverse a distance equal to its diameter (Eq. 4.2.1). The contact and compression stage is clearly visible on this plot, for within a time of about  $2\tau$  the projectile's initial kinetic energy is distributed among the four energy reservoirs in proportions derived in Section 4.5.2.

Subsequent to the contact and compression stage, the energy content of the various reservoirs evolves more slowly. Figure 5.5 shows that the internal energy in the target, which is by far the largest sink of energy, steadily increases until it accounts for nearly 90 percent of the total energy. The kinetic energy of the target, which is connected with the developing excavation flow, dec-



**Fig. 5.5** Partition of energy as a function of time in the vertical impact of a 10-km diameter anorthosite projectile on an anorthosite target at 30 km/second. All of the energy is initially kinetic energy of the projectile. However, after one or two  $\tau = L/v_i$  the energy is largely transferred to the target. After O'Keefe and Ahrens (1982a).

lines steadily after its initial peak near the end of contact and compression. This decline is initially caused by the decay of the high particle velocities in the expanding shock wave but continues later as the speed of the excavation flow itself declines. The internal and kinetic energies in the projectile become a nearly constant and insignificant proportion of the total energy after about  $10\tau$ . The lion's share of the projectile's initial energy is thus transferred to the target.

Approximate estimates of energy partition in an impact can be made using the Gault-Heitowitz or other approaches described in the last section. Such approximate models are useful so long as the expansion of the shock wave is reasonably simple. However, if it is desired to investigate energy partition into, say, a target planet's atmosphere, or ocean, or layers within the target, then numerical techniques are the only reasonable recourse.

### 5.2.6 The end of stress wave expansion: elastic radiation and seismic equivalents

As the strong shock wave generated by a hyper-velocity impact expands, it weakens, first degrading into a plastic stress wave preceded by an elastic precursor, then dying away into an entirely elastic wave. This elastic wave continues to travel through the target planet, further weakened only by normal seismic attenuation. The structure of the elastic wave generated by an impact is complex, with multiple pulses related to the elastic precursor, the elastic remnant of the plastic wave, and further reverberations resulting from free surfaces and interfaces with different seismic imped-

ance in the target. Surface waves, which attenuate less rapidly than body waves, are also excited by the impact and may be important at large distances.

Relatively little is known about the final elastic waves radiated by a large impact. A study of several small-scale missile impacts (Latham et al., 1970) suggests that their seismic radiation is similar in many respects to that of explosions, producing ground accelerations two to three times larger than surface explosions of the same energy release, but accelerations only 30 percent to 40 percent as large as buried explosions. Impacts are evidently more efficient in exciting *P*-waves than surface explosions, but not as efficient as buried explosions.

The *seismic efficiency*,  $k$ , relates the elastic energy ultimately radiated to the energy of the impact itself. It is defined by

$$E_{\text{seismic}} = kW \quad (5.2.14)$$

where  $E_{\text{seismic}}$  is the total elastic energy radiated and  $W$  is the initial kinetic energy of the projectile. Estimates of this efficiency range from  $10^{-3}$  to  $10^{-5}$ , with the most commonly accepted figure being  $k = 10^{-4}$  (Schultz and Gault, 1975). The amount of seismic energy radiated from a crater of a given diameter may be estimated with the help of the scaling relations described in Chapter 7, which relate the crater diameter to the energy and velocity of the impactor.

The seismic magnitude  $M$  produced by an impact can be estimated from the classic Gutenberg-Richter magnitude-energy relation,  $\log_{10} E_{\text{seismic}}$  (Joules) =  $4.8 + 1.5M$ . Inserting the seismic efficiency  $k = 10^{-4}$ , the relation between seismic magnitude and impact energy  $W$  is

$$M = 0.67 \log_{10} W - 5.87 \quad (5.2.15)$$

By this formula, a 30-m diameter projectile striking at 20 km/second would produce a seismic disturbance equivalent to a magnitude 4.8 earthquake. Using Equation 7.8.4, such an impact would produce a 1-km-diameter crater on the earth, about the size of Meteor Crater, Arizona.

The destructive effects of impact-generated seismic waves are not expected to be as severe as those of an earthquake of the same magnitude because an impact generates mostly *P*-waves, whereas an earthquake generates more destructive *S*-waves. A rule of thumb developed from explosion cratering experience suggests that an impact-generated seismic disturbance is equal in destructiveness to an earthquake one magnitude smaller.

Caution should be used in comparing the seismic effects of impacts and earthquakes. Not only are the types of waves emitted different, but the frequency spectra and duration of shaking may differ. Nevertheless, the difficulty experienced in

distinguishing earthquakes and explosions at teleseismic distances suggests that this comparison may have some validity.

The seismic energy emitted by large impacts may play a role in initiating mass movement and modifying the surface adjacent to the impact. The largest impacts have apparently produced distinctive terrains at their antipodes as a result of the convergence of seismic waves there. These topics are discussed in more detail in Section 12.6.

### 5.3 VAPOR PLUME EXPANSION

Some of the earliest and fastest "ejecta" to leave the site of an impact crater are not usually considered to be part of the excavation flow, although important amounts of material may be involved. This early flow is due to vaporized projectile and target material that expands out of the growing crater at high speed. Vapor plume expansion is not a familiar aspect of impact cratering in either laboratory scale experiments or missile impacts because the impact velocity in these circumstances is generally too low for vaporization to occur. Impact velocities must exceed about 10 km/second for significant amounts of vaporization in either silicate or water ice impactors or targets. Several numerical simulations of high velocity impacts have produced well-developed vapor plumes, although little attention has been paid to the phenomenon.

Recent studies have directed more attention to the vapor plume because of increasing interest in impacts on icy bodies in the outer solar system, the Cretaceous-Tertiary impact that evidently scattered iridium-rich dust over large areas of the earth, presumably by means of its expanding vapor plume, and because the vapor plume has been suggested as an agent for accelerating impact crater ejecta to high speed, even to Martian escape velocity, to explain how the SNC (Shergottite, Nakhelite, and Chassigny) meteorites might have left Mars' surface. Still more recent work suggests that the vapor plume ejected during an oblique impact between the protoearth and a Mars-size protoplanet might have created the earth's moon.

Although the vapor plume has not been much discussed in the cratering literature published to date, there is nothing intrinsically difficult about treating it numerically. Future work will, it is hoped, rectify this omission. In the meantime, I will present an order-of-magnitude description of vapor plume expansion based on simple analytic models. Readers requiring a more detailed study of vapor plume expansion are advised to find themselves a computer and a hydrocode and go to work!

Vapor plume expansion begins as soon as the rarefaction from the rear of the projectile releases it from high pressure, during the contact and compression stage (see Sec. 4.5.3). If the unloaded

projectile and target are in the vapor phase, they begin to expand upward and outward at high speed. The expanding gas is, of course, a mixture of components shocked to different pressures and thus with different initial temperatures. The expansion is dominated by the hottest gases, however, and for the sake of simplicity I will assume that the gas starts out in a uniform thermodynamic state.

Numerical computations show that the vapor expands out of the crater just after the first ejecta have been thrown out from the target's surface adjacent to the impact. Some computations show vapor expansion beginning in an annulus surrounding the impactor, where decompression of the shocked projectile and target begins. The rear of the projectile starts to decompress somewhat later, after it has penetrated deeper into the target. This produces a rather complicated initial flow pattern illustrated schematically in Figure 5.6a. Some *inward*-directed velocity components even develop as the hot gas from the rim of the growing crater expands back over its center.

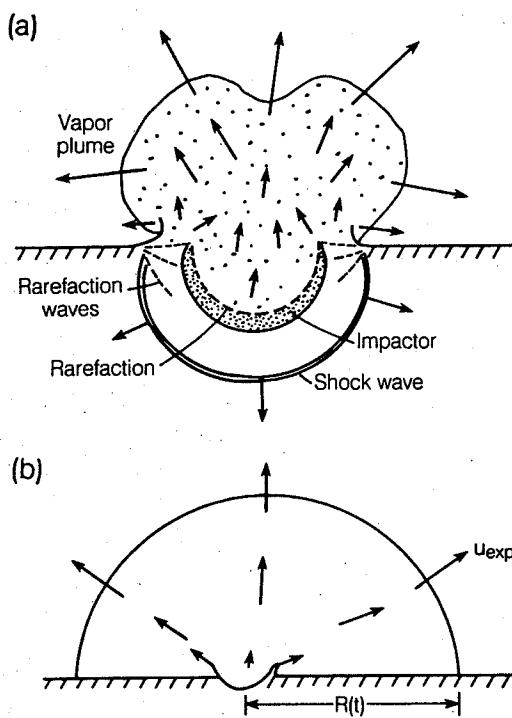
The effects of oblique impact on vapor plume expansion are not well understood. The initial high velocity vapor jet was discussed in Section 4.3, but this jet is a special phenomenon produced by the oblique convergence of shock waves from the projectile and target and the expansion of the bulk of the vapor produced later was not addressed. Perhaps the simplest procedure is to add a horizontal velocity component to the expanding gas cloud in the direction of the projectile's motion, but it is not clear how large this velocity should be. More studies of oblique impacts are urgently needed.

Once the vapor plume has expanded to several times the projectile diameter, the flow field becomes simpler and can be approximated as an expanding hemisphere of hot gas, shown in Figure 5.6b. The gas plume tends to outrun the classic ejecta curtain, enclosing the developing crater in a transient atmosphere of vaporized projectile and target. The plume continues to accelerate, driven by the pressure gradient between its interior and the surrounding vacuum (the special behavior of a vapor plume that expands into a surrounding atmosphere is treated in Section 11.3). The acceleration,  $a(r)$ , of a parcel of gas at radius  $r$  in a hemispherically expanding gas cloud is given by

$$\ddot{a}(r) = - \frac{1}{\rho_{\text{gas}}} \frac{dP}{dr} \quad (5.3.1)$$

where  $\rho_{\text{gas}}$  is the gas's density. For our purposes it is adequate to treat the gas as perfect, with an equation of state

$$P = \frac{\rho_{\text{gas}} R_g T}{\mu} = (\gamma - 1) \rho_{\text{gas}} E_{\text{gas}} \quad (5.3.2)$$



**Fig. 5.6** Expansion of vapor produced during the early stages of an impact. Part (a) shows the initial stages of the expansion where the vapor flow pattern may be complex as a result of the timing of release of various portions of the projectile and target from high pressure. Pressure release first occurs near the rim of the growing crater so that initial vapor expansion occurs there. The remainder of the projectile and target beneath it begins expanding later, producing a dominantly upward flow in the center of the crater, although small inward velocity components may occur at this stage. Part (b) illustrates the flow somewhat later, when the expanding vapor plume is nearly hemispherical and is moving so rapidly that its front outdistances the solid ejecta curtain. The expansion velocity in the vapor plume increases nearly linearly with distance from the crater's center to a maximum  $u_{\text{exp}}$  at the plume's edge.

where  $R_g$  is the gas constant, 8.317 J/mole K,  $T$  is the gas temperature,  $E_{\text{gas}}$  is its specific internal energy,  $\mu$  is its kilogram-equivalent mean molecular weight, and  $\gamma$  is the ratio of its specific heats. For silicate vapors up to about 15,000 K, detailed equation of state modeling shows that  $\gamma = \frac{5}{3}$  and  $\mu = 2 \times 10^{-2}$  kg/mole yields an adequate approximation right up to the liquid-vapor phase boundary.

Elementary thermodynamics shows that the mean velocity of the expanding gas ultimately reaches  $u_{\infty}$ , where

$$u_{\infty} = \sqrt{2E_0} \quad (5.3.3)$$

$E_0$  is the initial specific energy of the gas. It can be roughly taken to be  $E_0 = (E_p - H_{\text{vap}})$  where  $E_p$  is the specific internal energy of the projectile and  $E_t$  is the same quantity for the target material, whichever is appropriate. These specific energies can be estimated from the planar impact approximation, Section 4.5.1.  $H_{\text{vap}}$  is the enthalpy of vaporization of the material, typically about 13 MJ/kg for silicates.

The expansion of the vapor cloud can be followed in more detail by a simple dynamical model described by Zel'dovich and Raizer (1967, p. 103). In this model the expansion velocity in the hemispherical gas cloud increases linearly from zero at its center to  $u_{\text{exp}}$  at its outer edge, located at radius  $R(t)$ . This maximum expansion velocity is given by

$$u_{\text{exp}} = [2\gamma/(\gamma - 1)]^{1/2} u_{\infty} \quad (5.3.4)$$

The expansion velocity  $u(r)$  at any given radius  $r$  is simply  $u(r) = (r/R)u_{\text{exp}}$ . The density  $\rho_{\text{gas}}$  in the cloud, which is uniform, declines from its initial value  $\rho_0$  as the cloud expands from its initial radius  $R_0$

$$\rho_{\text{gas}} = \rho_0[R_0/R(t)]^3 \quad (5.3.5)$$

$R_0$  is comparable to the diameter of the projectile. The gas pressure is a roughly parabolic function of distance from the center,

$$P(r) = P_c[1 - (r/R(t))^2] \quad (5.3.6)$$

where  $P_c$  is the central pressure of the gas cloud.  $P_c$  itself declines from its initial value  $P_0$  as the cloud expands,

$$P_c = P_0[R_0/R(t)]^{3\gamma} \quad (5.3.7)$$

Temperatures at any desired radius and time may be deduced from the pressure, Equation 5.3.6 and the density, Equation 5.3.5, using the perfect gas equation of state, Equation 5.3.2. For any given parcel of gas within the cloud,  $T$  declines adiabatically as  $R(t)^{-3(\gamma-1)}$ . Finally, the radius  $R(t)$  of the gas cloud increases approximately as

$$R(t) = [(u_{\text{exp}} t)^2 + R_0^2]^{1/2} \quad (5.3.8)$$

Although Equations 5.3.5 through 5.3.8 are usually approximate, it can be shown that in the special case  $\gamma = \frac{5}{3}$  they become an exact solution to the equations of motion of a hemispherically expanding gas cloud, starting from a uniform initial density and a parabolic pressure distribution, Equation 5.3.6.

These analytic solutions show that the gas cloud expands with enormous speed. The edge of the gas cloud moves about three times faster than the mean gas velocity, which itself may exceed 10 km/second in a 20- to 30-km/second impact. The gas blowing out of the crater has an initial density only slightly reduced from the target's initial density, yet it moves at speeds in excess of 30 km/second. Even after it has expanded to a diameter

equal to that of the eventual crater, its density and velocity are sufficient to accelerate the solid ejecta appreciably, most of which are thrown out only after the gas cloud has swept over the target's surface (Vickery, 1986). The acceleration of the gas cloud is so violent that large ejecta fragments may be crushed by aerodynamic pressures generated by the differential velocity between the expanding vapor plume and ejecta thrown out from the growing crater.

As the hot vapor plume expands, it eventually begins to condense. Conditions for condensation into either liquid or solid particles are generally first reached in the cloud's center and then spread outward, although in free expansion of this type up to 50 percent of the vapor may never condense, ending up as free atoms or molecules in the vacuum of space. Once condensation begins, the perfect gas law no longer describes the relation between pressure and temperature in the cloud; hence the approximate dynamical model described above ceases to hold. The expansion of a two-phase gas cloud can only be treated numerically. The equation of state of the two-phase mixture is, however, quite simple. Once the phase curve is reached from the vapor side, pressures and temperatures remain on the curve. For a silicate vapor with a composition similar to that of Forsterite ( $Mg_2SiO_4$ ), a good representation of the phase curve has been found to be

$$P = 2.2 \times 10^5 e^{-48,900/T} \quad (5.3.9)$$

Where the pressure  $P$  is in MPa and the temperature  $T$  is in Kelvin. After condensation begins, pressure and temperature are linked by this equation for the remainder of the expansion. If the expanding vapor first reaches the phase curve at pressure  $P_1$  and temperature  $T_1$ , then the mass fraction of condensed gas  $X_c$  is

$$X_c = [R_g \ln(P_1/P) - C_{pg} \ln(T_1/T)] / (H_{vap}/T) \quad (5.3.10)$$

$C_{pg}$  is the specific heat at constant pressure of the gas phase, about 27 J/mole K for Forsterite vapor, and  $H_{vap}$  is the enthalpy of vaporization, about 640 kJ/mole for Forsterite.

It would be useful to be able to estimate the size of the condensed liquid droplets or solid dust particles from the expanding vapor cloud. Unfortunately, there is little experimental information on this topic. One estimate (Raizer, 1960), derived for the condensation of iron vapor formed by an impact, assumes the validity of homogeneous nucleation theory. A similar estimate for silicate vapors would probably not be much different. A useful form of Raizer's result (O'Keefe and Ahrens, 1982a) is

$$d_{\text{drop}} = 2.48 \times 10^{-4} L v_i^{2/3} (E_0 / 8.5 \times 10^6)^{-5.8} \quad (5.3.11)$$

where  $d_{\text{drop}}$  is the diameter of the condensate particle in m,  $L$  is the projectile diameter in m,  $v_i$  is its impact velocity in m/second, and  $E_0$  is the initial specific internal energy of the expanding gas in J/kg. This equation predicts that the vapor plume from a 10-km diameter projectile striking at 25 km/second would condense into particles roughly 400  $\mu\text{m}$  in diameter.

Unfortunately for Equation 5.3.11, recent experiments on the condensation of metal and silicate vapors seem to indicate that classical homogeneous nucleation theory fails to describe these materials. The critical-size condensation nucleus consists of only a few atoms, and so the density of initial nucleation sites is extremely large, leading to condensates that are themselves only a few microns in diameter. Further work is obviously necessary before the size of the condensed particles can be stated with any degree of confidence. This issue is of critical importance to the theory that the Cretaceous era on earth was ended by the impact of a 10-km diameter comet or asteroid, since the dust particles produced by such an impact would have to be smaller than a few microns in diameter for them to darken earth's skies for an appreciable period. Particles as large as 100 microns fall out in less than a day and so could not contribute to the extinction of terrestrial lifeforms.

The only real data on vapor plume condensation may be derived from the presence of large numbers of 100 to 200  $\mu\text{m}$  nickel-iron spherules around Meteor Crater, Arizona. They were discovered by H. H. Nininger in 1946. Nininger believes that they condensed from the vaporized nickel-iron projectile that produced the crater, and their abundance supports his view (Nininger, 1956). However, the projectile that formed Meteor Crater was only on the order of 30 m in diameter, so that the spherules are even larger than Equation 5.3.11 predicts. Either the presence of earth's atmosphere slowed the vapor cloud expansion and so allowed the nickel-iron droplets to grow to larger size than Equation 5.3.11 predicts (this equation was derived for free expansion into a vacuum) or perhaps heterogeneous nucleation dominates the condensation. In this latter case the size of the condensates cannot be derived without an estimate of the number density of nucleation centers. Another, completely different, explanation is that the spherules originally formed from splashes of melted, but not vaporized, nickel-iron. In this case their size has little bearing on the vapor condensate issue.

Besides droplets condensed from the expanding vapor, the vapor plume generated by an impact may also carry up droplets of impact melt ejected from the crater during vapor plume expansion. Tektites, glassy blobs of once-molten rock found strewn widely across the earth's surface, may have

originated in this way. The glass spheres commonly found in lunar soils may similarly have originated in impact events. The *minimum*-size droplet that can be produced from a melt originates from the breakup of tiny filaments that squirt out from the moving liquid. Its diameter  $d_{\min}$  is determined by a balance between the surface tension  $\sigma$  of the melt, which tends to hold small parcels of melt together, and the velocity of the filament,  $v_{\text{jet}}$ , which tends to disperse the liquid:

$$d_{\min} = 8\sigma/(\rho_{\text{melt}} v_{\text{jet}}^2) \quad (5.3.12)$$

where  $\rho_{\text{melt}}$  is the melt density. The jet velocity is at most equal to the impact velocity,  $v_i$ . For a typical silicate the surface tension  $\sigma$  is about 0.3 N/m, so for a filament velocity on the order of 1 km/second the minimum droplet size is about 1 nm.

The *maximum*-size droplet that is stable in the ejecta plume is determined by a balance between surface tension forces and aerodynamic forces, which tend to disrupt the droplet. This diameter  $d_{\max}$  is given by

$$d_{\max} = 10\sigma/[\rho_{\text{gas}}(\Delta v)^2] \quad (5.3.13)$$

where  $\rho_{\text{gas}}$  is the ambient density of the gas in the vapor plume and  $\Delta v$  is the differential velocity between the melt droplet and the gas.

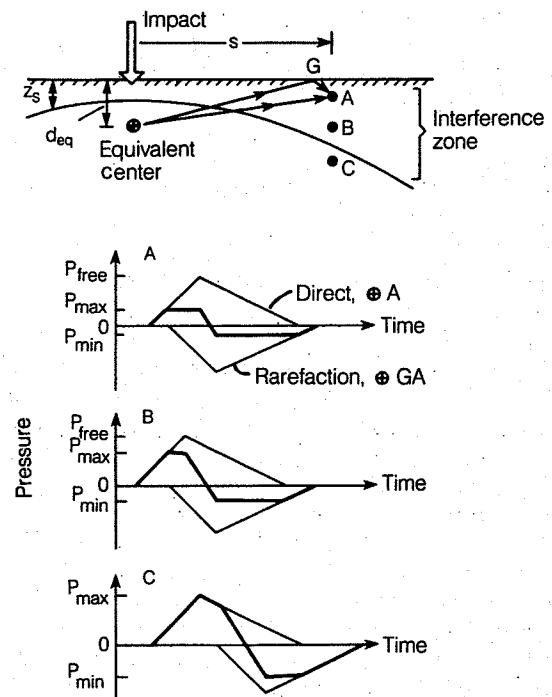
Many other aspects of the expansion of the vapor plume produced by an impact have yet to be investigated. Study of the geological and physical effects of this phenomenon may yield unexpected rewards, in addition to an understanding of the origin of tektites or of the details of the terminal Cretaceous event.

#### 5.4 THE NEAR-SURFACE ZONE

As the shock wave expands hemispherically away from the site of the impact, shallow portions of it begin to interact with the target's free surface. The geometry of the expanding shock is that of a hemisphere centered on a point some distance below the surface, here called the *equivalent center*, so that the strong shock waves generated by the impact impinge on the surface from below. Since the free surface is, by definition, a plane of zero pressure, a physical dilemma is created as the rising shock wave, with very high pressures behind it, meets the surface.

This dilemma is resolved by a rarefaction wave, equal in strength but of opposite sign to the shock wave, which starts downward from the surface as soon as the shock wave arrives. The sum of the pressure in the two waves is always zero on the free surface, as required physically, but the situation is more complex below the surface. Points a short distance below the surface are subjected first to compression, then to tension as the rarefaction wave makes its presence felt, as shown in Figure

5.7. The shock is approximated as a triangular pulse with rise time  $\tau$ , (Eq. 5.2.1), followed by a longer decaying tail. At near-surface point A in Figure 5.7, the rarefaction arrives before the shock has reached its peak, reducing the maximum pressure at this depth to a fraction of the free shock pressure (the maximum shock pressure elsewhere in the target at the same distance from the equivalent center). Similar events occur at point B in Figure 5.7, although because of its



**Fig. 5.7** Stress-wave interference near the free surface of the target. The stress waves appear to radiate from a center at depth  $d_{\text{eq}}$  below the target surface. At point A the direct wave, represented by a triangular pulse with a short rise time and longer decay time, arrives first, but is shortly followed by a tensile rarefaction wave. The rarefaction is produced by the reflection of a compressional wave at G. The direct and rarefaction waves interfere at A to limit the net pressure excursion  $P_{\max}$ , shown by the heavy line, to a fraction of the pressure excursion in the absence of interference,  $P_{\text{free}}$ . A later tensile phase may fracture the target rock, producing spalls and Grady-Kipp fragments. At the deeper point B the rarefaction arrives later and the pressure reduction by wave interference is smaller. At point C, below the interference zone, the rarefaction arrives after the direct wave has reached its peak and no pressure reduction occurs. Note, however, that a strong tensile phase still develops, so that tensile rock fracture extends well beneath the interference zone.

greater depth, the maximum pressure is a larger fraction of the free shock pressure. Point C is so deep that the rarefaction does not arrive until after the shock wave has reached its peak and so no reduction of the maximum pressure occurs, although the rarefaction eventually makes itself felt.

The near-surface zone in which the actual maximum pressure is reduced from the free maximum pressure by interference between a rarefaction and the direct shock wave is the *interference zone*. Its depth  $z_s$  is a function of the shock wave rise time  $\tau$ , and the local wave velocity  $U_r$ . Differences in the shock and rarefaction wave velocities, although potentially important, are ignored here for simplicity. The actual shock structure may become very complex, but the ultimate results are represented reasonably well by the picture described here. The depth of the bottom of the interference zone is approximately given by

$$z_s = (U_r \tau / 2) \{1 + s^2 / [d_{eq}^2 + (U_r \tau / 2)^2]\}^{1/2} \quad (5.4.1)$$

where  $s$  is the horizontal distance across the surface from the impact site to the position where the depth of the interference zone is evaluated.

The product  $U_r \tau$  is the width of the rising portion of the shock wave. It is roughly equal to the projectile's radius. In older work this width was approximated as zero. In this case the interference zone vanishes,  $z_s = 0$ , and the interesting phenomena that occur in this zone were missed.

The pressure profiles at the three depths A, B, and C not only show the reduction of the maximum compression in the interference zone, but also shown is a later *tensional* phase due to the rarefaction. In Figure 5.7, where tensional failure of the target is not permitted, this tension may reach a maximum comparable in magnitude to the maximum compressional stress attained at that depth. Tensional stresses of many GPa are clearly unphysical because they greatly exceed the target's intrinsic tensile strength. Well before such a stress is reached, cracks nucleate in the target and lengthen at speeds approaching the speed of sound. As these cracks link up, the target material loses cohesion and disaggregates into a mass of broken fragments whose size can be roughly estimated by Equation 6.4.2. The tensional stress excursion is thus halted by the fragmentation of the target and is held to a maximum value equal to the target's dynamic tensile strength, typically some 100s of MPa for most rock materials.

This tensile fragmentation of the target begins some distance below the surface, where the maximum tensional stress in the interfering stress waves first reaches the dynamic tensile strength. Rock lying at shallower depths is thus not broken, but instead is thrown off the surface at high velocity in the form of thin *spall plates*. These are the largest and least shocked fragments thrown out at any given velocity. The thickness of these spall plates is given by Equation 6.4.1. Only one spall

zone forms at the surface of homogeneous targets, although layering may produce several. At high ejection velocities the spall plates contain so much elastic energy that they tend to break up into smaller fragments. Although spall plates represent only a small amount of material, they are probably responsible for the secondary crater fields observed around large impact craters. See Section 6.4 for more discussion of this phenomenon.

The tensional phase of the rarefaction is not, however, limited to the near-surface interference zone. At depth C in Figure 5.7, below the interference zone, the tensional stress excursion is larger than at any of the other points illustrated. The rarefaction wave that originates at the free surface thus propagates downward to great depths, fracturing the target rock as it goes. This is the origin of the brecciation and fracturing detected by geophysical means beneath both terrestrial and lunar impact craters. Contrary to the common notion, the rock beneath an impact crater is not "crushed" by the shock wave: this requires stresses in excess of the Hugoniot Elastic Limit (Sec. 3.3.3), which are not attained at any significant distance away from the crater's center. Instead, the rarefaction following the shock propagates downward and outward many times the crater depth or diameter, fracturing the rock *in tension* as it goes.

Stress-wave interference not only shatters the near-surface rocks, it also ejects them at very high speeds. High ejection speeds develop in spite of the low pressures in the interference zone because acceleration depends on the pressure gradient, not the pressure itself. As discussed in connection with Equation 5.3.1, material near the target surface is subject to an acceleration

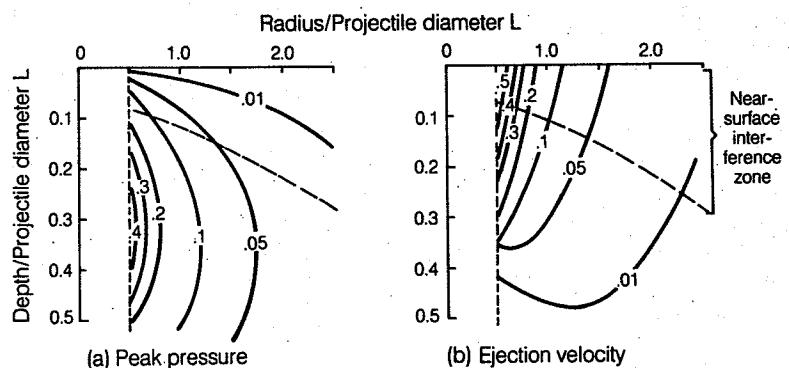
$$a = \frac{1}{\rho_t} \frac{dP}{dz} \quad (5.3.2)$$

where  $\rho_t$  is the target's density and  $z$  is the vertical coordinate. Although the pressure is zero at the free surface, interference with the rarefaction actually doubles the pressure gradient for a vertically impinging shock wave. In this case, target material ejected from the free surface thus leaves at *twice* the particle velocity of the undisturbed shock wave.

The velocity increase upon ejection is not as dramatic for a shock wave impinging obliquely upon a free surface, but it is still substantially larger than the velocity imparted by the subsequent excavation flow. In a simple model of stress wave interference (Melosh, 1984), the ejection velocity  $v_e$  is

$$v_e \approx 2u(r)[1 + (s/d_{eq})^2]^{-1/2} \quad (5.3.3)$$

where  $u(r)$  is the particle velocity in the undisturbed shock at a distance  $r = (s^2 + d_{eq}^2)^{1/2}$  from the equivalent center. The ejection velocity dec-



**Fig. 5.8** Contours of peak pressure and ejection velocity due to stress-wave interference near the surface of the target from a simple model (Melosh, 1987). Note the  $5\times$  vertical exaggeration necessary to show the details. The pressure is in units of  $\rho_0 v_i^2$  and the velocity is in units of  $v_i$ , the impact velocity. The pressure increases rapidly with depth beneath the surface but the ejection velocity, high near the surface, declines less rapidly, resulting in a significant volume of moderately shocked target material ejected at high speed. Note that this ejection velocity does *not* include the additional component related to the excavation flow: in reality, rather than falling to zero at depths greater than about half a projectile diameter, the velocity falls to the excavation flow velocity, which is about 20 percent of the surface velocity as a result of stress wave interference.

lines with increasing depth in the near-surface zone, shown in Figure 5.8, which illustrates both ejection velocity (the portion of it due to stress wave interference) and pressure at  $5\times$  vertical exaggeration.

Not only is near-surface material ejected at high velocity, but it is also ejected early compared with the main body of the ejecta. Because its velocity is a result of stress wave interference, this material is ejected just after the shock wave arrives. Its ejection time is thus a nearly linear function of distance away from the impact site. Material thrown out by the excavation flow leaves at a time proportional to the square or higher power of the distance from the impact site.

The near-surface zone is thus the source of an extraordinary body of ejecta. Because of the presence of a free surface, this material is lightly shocked compared with deeper-lying material at a comparable distance from the impact. It is ejected at the highest speed of any solid ejecta, speeds that reach a maximum of about half the impact velocity of the projectile, although Equation 5.3.3 shows that this velocity falls roughly as  $1/s^3$  at large distances from the impact site. It is also ejected early—it clearly leads the main body of ejecta and is preceded only by jettied material and perhaps by the vapor cloud. The total volume of this extraordinary material, however, is small. The thickness  $z_s$  of the interference zone is comparable to the projectile diameter, which is much smaller than the crater diameter or depth. At most, this zone encompasses only about 10 percent of the total ejecta, and the fastest, most lightly shocked ejecta originating nearest the target surface may be only a small percentage of the total. The volume of material ejected to a given speed and with a shock level less than some given

maximum can be estimated from Equation 6.4.3 in Section 6.4, where a more detailed discussion of the ejecta from this zone is presented.

Ejection of fast, lightly shocked near-surface material by stress wave interference probably accounts for the recently discovered meteorites of lunar origin. These show no evidence of shock levels greater than about 15 GPa, in spite of their minimum ejection velocity of 2.5 km/second. Cosmic-ray-produced nuclei in these meteorites show that they were exposed at the lunar surface prior to their ejection, in agreement with the stress wave interference model. A similar mechanism might account for the ejection of the SNC meteorites from the near-surface zone of a large crater on Mars. Many other meteorites seem to have originated from near the surface of their parent bodies, making the existence of, and interactions in, the interference zone of greater scientific interest than the fraction of the material involved would suggest.

An interesting inverse application of stress wave interference near a free surface is suggested by several millimeter-diameter fragments of unshocked chondrite meteorite found in lunar soil samples. How they reached the lunar surface without suffering shock upon impact has been considered a puzzle. However, the rear of the projectile in an impact is also a free surface, and because the shock wave traversing the projectile also has a finite rise time, stress wave interference there would eject a small fraction of its mass at high velocity *relative to the projectile*. This material may have a small velocity relative to the target, however, and unshocked projectile fragments may thus be deposited gently on the target planet's surface.

Although the free surfaces of either the target or

projectile are ordinarily the sites of the largest impedance contrast and hence develop the strongest reflections, subsidiary reflections may develop on the boundaries of layers in the target with smaller impedance contrasts. The physics of such reflections is discussed briefly in Section 3.3.2 for normal incidence of an elastic wave on an interface. The result of these stress wave reflections and subsequent interactions would be a very complex structure in the expanding shock wave that influences petrographic shock indicators and gives a geologic picture of highly variable shock levels at a given range from the impact. Such interactions are not well understood or even much investigated at the present time.

## 5.5 THE EXCAVATION FLOW

### 5.5.1 Initiation

The rapid passage of the shock wave through the target leaves the material behind it in motion. At any given distance from the impact site the material's initial velocity is between one-third and one-fifth of the peak particle velocity in the shock at that distance, depending upon the equation of state of the target material (see Section 3.4.3 for the origin of this residual velocity). This velocity is directed radially away from the impact site immediately behind the shock wave. However, rarefaction waves propagating downward from the free surface create an upward-directed pressure gradient behind the shock, which adds an upward component to the radial velocity, ultimately producing the upward-and-outward excavation flow pattern illustrated schematically in Figure 5.1. Material located directly beneath the impact site continues moving downward and is never ejected from the growing crater. The flow field thus left behind by the shock wave is called the excavation flow. It is principally responsible for opening the crater.

The expanding shock wave moves so much faster than the excavation flow that the two can be treated separately for most purposes: by the time the excavation flow is fully developed and the crater begins to grow, the shock wave is long gone and does not need to be considered any further. The displacements produced in the target by the shock are also negligible (they are at most equal to a projectile diameter at the impact site and fall off rapidly at greater distances). Nearly all of the displacements in an impact crater develop during the excavation flow.

Excavation flow velocities are highest near the impact site and fall off approximately as an inverse power of the distance from it. Since even the highest velocities are between one-third and one-fifth of the peak particle velocity in the shock, which itself is typically one-half of the impact velocity, the maximum excavation flow velocity is

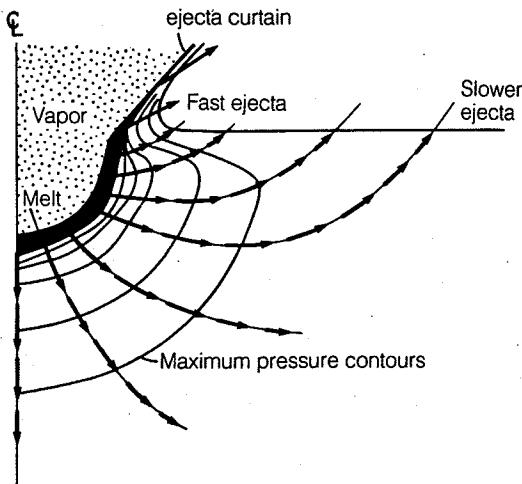
between one-sixth and one-tenth of the impact velocity. Even at impact velocities as high as 40 km/second, the excavation flow is never faster than about 6 km/second—less than the speed of sound in most rock material. This maximum velocity declines rapidly with both time after the impact and distance away from it. The excavation flow is thus predominantly subsonic and hence can be treated as incompressible in all but the highest-velocity planetary impacts.

A common, but incorrect, idea is that the excavation flow is somehow driven by the projectile as it continues to plunge into the target. This notion is probably derived from observations of low-velocity impacts in which the projectile's momentum is coupled into the target so slowly that it actually does drive the flow. However, in hypervelocity impacts the projectile's initial energy and momentum are almost entirely coupled into the target during the brief contact and compression stage, after which the projectile plays little role in the developing flow. In high-speed impacts, the projectile and part of the adjacent target actually vaporize and, far from continuing to push material out of the crater, the projectile expands back out of it as a hot gas.

### 5.5.2 Geometry of the excavation flow

The excavation flow field created by the shock wave's passage possesses a characteristic form, which is illustrated schematically in Figure 5.9. Velocity vectors in the material surrounding the growing crater can be connected by curved lines called *streamlines*, which begin on the inner surface of the growing crater and progress outward, following paths that are parallel to the velocity vectors at each point. The streamlines near the lip of the crater curve upward and outward. Those originating lower on the growing crater's wall dip down before turning upward, following material ejected from the crater. The streamlines that originate on the crater's floor never reach the surface, but instead indicate the paths of target material that is displaced downward. Space to accommodate this material is created by plastic deformation of the surrounding rocks.

Material in the excavation flow is considered to be ejected when it rises above the preimpact surface. At this stage in the diverging excavation flow, portions of the moving debris lose contact with their neighbors and begin to follow purely ballistic trajectories. The ejecta curtain, Figure 5.10, takes on the form of an inverted cone whose lower edge defines the outer lip of the growing crater. Ejecta fragments in this curtain seldom interact with one another: the apparently coherent form is due to a coincidence in alignment of many fragments, each following its own free trajectory. The detailed form of the ejecta curtain is controlled by the velocity, angle, and time of ejection



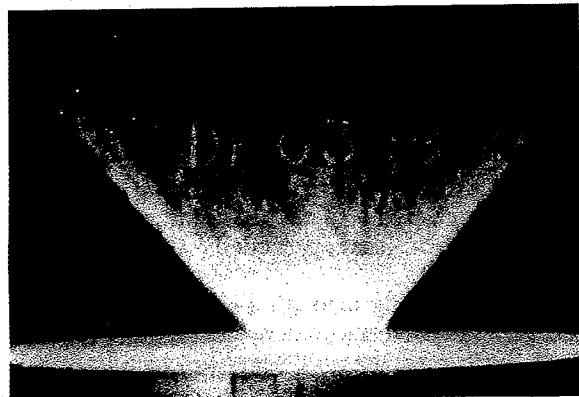
**Fig. 5.9** Geometry of the excavation flow field. The arrows illustrate the upward and outward excavation flow left behind by the rapidly expanding shock wave, which has moved beyond the boundaries of this illustration. The velocity vectors are connected by streamlines that define the boundaries of streamtubes. In a steady flow, material within each streamtube stays within that tube as it expands. Also shown are contours of peak shock pressure. It is clear that streamlines cut across pressure contours, so that material ejected from a given streamtube contains a mixture of shock levels. Even the lowest velocity ejecta will contain some highly shocked impact melt. When a material following a streamline crosses the initial surface it is ejected ballistically and forms part of the ejecta curtain. Ejecta from near the impact site travels at high speed, whereas ejecta emerging at larger distances travels at slower velocities.

of each of its component particles. In most small-scale experiments that have been performed in a vacuum, however, the ejecta curtain is a nearly straight-sided inverted cone whose sides make an angle of about 45 degrees with the target surface.

Since the physics of the ejecta curtain's motion is very different from that of the excavation flow, further discussion of the ejecta curtain is deferred until Chapter 6 (Sec. 6.3.3) where we will see that its dynamics plays an important role in the deposits produced by the ejecta through the process of ballistic sedimentation. The time and velocity of ejection are discussed in Section 7.12. The remainder of this chapter will deal with the nature of the excavation flow only up until the point where material is ejected.

Returning to Figure 5.9, a very important feature of the excavation flow is the fact that the streamlines cut across contours of maximum shock pressure. Material ultimately ejected from the crater along a given streamline thus contains contributions from all of the different shock levels

intersected. In general, the highest-velocity ejecta are the most highly shocked, but ejecta thrown out at any given velocity always contain some material shocked to the highest levels. Even the very lowest-velocity ejecta, the material that eventually forms the crater rim, may contain a significant amount of highly shocked impact melt. Such melt veneers are often present on the rims of small lunar impact craters.



**Fig. 5.10** The ejecta curtain produced in a small-scale experiment forms an inverted cone that expands with time. *Photo courtesy of P. H. Schultz.*

## IMPACT CRATERING

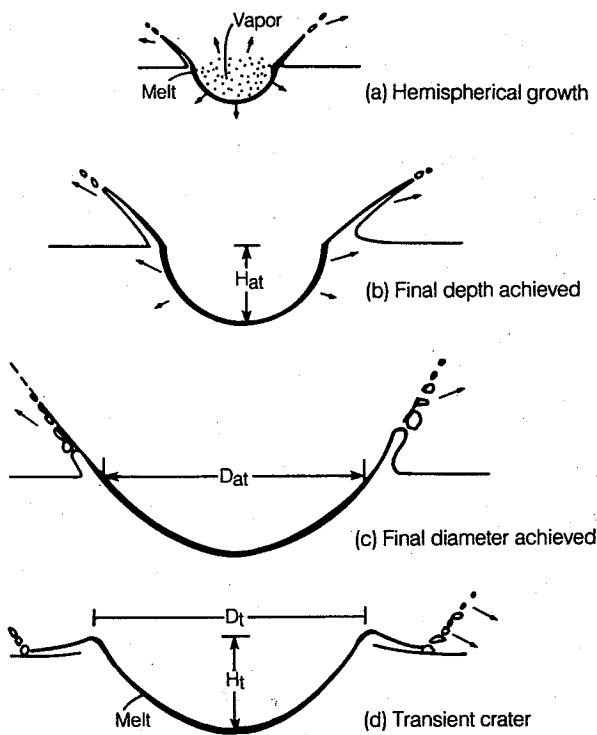
The streamlines illustrated in Figure 5.9 may be considered to form the bounding surfaces of *streamtubes* if the flow is steady (remember that the lines shown on the two-dimensional section in Figure 5.9 are really axially symmetric surfaces in space). Although the cratering flow is not strictly steady, it approaches a steady flow closely enough for the streamtube concept to be useful. Moving material remains within a given streamtube: none crosses from one streamtube into another because the transverse velocity component along a streamline is zero. The complex excavation flow can thus be broken down into the motion of material in each individual streamtube.

Since the excavation flow is nearly incompressible, Bernoulli's law applies to material flowing within a streamtube. The cross-sectional area of each streamtube increases rapidly as the radius increases, simply because of the axial symmetry of the flow, as well as the divergence of the streamlines in the section illustrated. The velocity of the material flowing through a streamtube thus decreases as it moves outward. Although highly shocked material from close to the impact site initially moves at high speed, it thus slows as it travels outward. Its final velocity upon ejection depends upon how far it has to travel before it reaches the surface. Ejecta emerging from short streamtubes travel at high speed, whereas ejecta emerging from long streamtubes travel at low speed. The small volume of highly shocked material on the inner surface of the crater spreads to cover the increasingly large area of the streamtube as it travels. When it emerges from the target surface it is ejected no faster than the more lightly shocked debris that preceded it through the streamtube.

Material flowing through a given streamtube is also generally decelerated by forces acting on it. Drag (both positive and negative) against material in adjacent streamtubes, forces due to internal deformation, and gravity all slow the flow down. Although these retarding forces are not sufficient to prevent the ejection of material in the upper streamtubes, material within deeper streamtubes emerges at progressively smaller velocities until eventually the ejection velocity declines to zero. The streamtube in which this occurs defines the inner rim of the transient crater. Material flowing in still deeper streamtubes is not ejected ballistically, but may merely uplift the surface outside the crater.

### 5.5.3 Crater growth

Experimental and numerical studies of impact crater formation show that the excavation flow originally opens an approximately hemispherical cavity (Fig. 5.11a). This cavity is first created during the contact and compression stage, when its



**Fig. 5.11** Growth of a crater. The crater becomes hemispherical a short time after the impact (a) and initially expands at a fraction of the impact velocity. Its rate of growth in depth slows and finally ceases (b) before its radial growth halts (c). The resulting crater (d) is called a transient crater because it is subject to further gravitational collapse, described in Chapter 8. At all stages of its expansion the crater is lined with melt or highly shocked rocks originally produced near the site of the impact.

diameter is comparable to the projectile's diameter. It is lined with impact melt or the most highly shocked material produced by the impact. The cavity grows at a steadily decreasing rate until its maximum depth is attained (Fig. 5.11b).

The shape of the crater cavity at any given time is determined by the position reached by the innermost material within any given streamtube. The maximum crater depth is thus equal to the depth attained by material in the streamtube that plunges down vertically beneath the center of the crater (often called the axial streamtube). Although gravity tends to accelerate this material, the strength of the target and increasing lithostatic pressure from the surrounding rock eventually halt its motion. In the absence of large enough strength, the material thus pushed downward might begin to rise after it achieves its maximum depth. Such rebound is observed in large craters, discussed in Section 8.3.

Neglecting strength, the time  $T_d$  required for the crater to grow to its maximum depth below the

preimpact surface,  $H_{at}$ , is roughly the time for an object to fall freely the same distance,

$$T_d \approx (2H_{at}/g)^{1/2} \quad (5.5.1)$$

where  $g$  is the surface acceleration of gravity. Pursuing the free fall analogy a little further, if the mean initial velocity in the streamtube is of order  $\langle u \rangle_i$ , then the crater's depth is  $H_{at} \approx \langle u \rangle_i^2/2g$ . Unfortunately, it is difficult to define a precise value for  $\langle u \rangle_i$  in the absence of a dynamical model for the cratering flow, although the Z-model discussed in Section 5.5.6 permits some estimates.

Although the crater's growth in depth is halted relatively early by the resistance of the underlying target rocks, there is less resistance to material moving through streamtubes that intersect the surface near the rim of the growing crater. The crater thus continues to grow in diameter after its growth in depth has halted. Excavation finally ceases when no more material is ejected from the crater and its final diameter  $D_{at}$  (the "apparent diameter" measured at the preimpact surface) is achieved (Fig. 5.11c). The resulting crater is thus broader than a hemisphere. It can be approximately described as a paraboloid of revolution.

The formation time  $T_f$  of a crater whose excavation is halted by gravity is discussed in Section 7.11. It is typically equal to the period of a gravity wave of length equal to the crater's diameter,

$$T_f \approx (D_{at}/g)^{1/2} \quad (5.5.2)$$

Note that for a hemispherical crater,  $T_f$  in Equation 5.5.2 and  $T_d$  in Equation 5.5.1 are identical.  $T_f$  is longer than  $T_d$  when  $D_{at}$  is larger than  $2H_{at}$ —

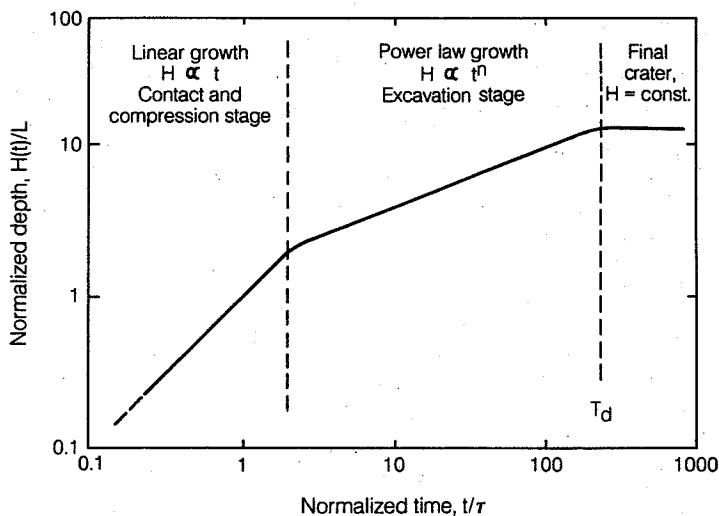
that is, when the crater is broader than a hemisphere.

A given parameter such as depth or diameter is observed to grow as a power of time during most of the excavation stage in both small-scale experimental impact craters and in numerical computations. This power-law growth is probably a consequence of the self-similarity of the excavation flow discussed in Section 7.12. Figure 5.12 illustrates the increase in crater depth as a function of time. The depth initially increases linearly with time during the contact and compression stage. Material on the crater floor moves downward at constant velocity. However, after a time equal to a few  $\tau$ , growth slows and the depth depends upon a power of  $t$ ,

$$H(t) \propto t^n \quad (5.5.3)$$

where  $n \approx 0.4$  in many experiments (Holsapple and Schmidt, 1987). Eventually, after a time given approximately by Equation 5.5.1, the crater's growth ceases and its depth takes on a constant value equal to the depth of the transient crater,  $H_{at}$ . The crater's diameter grows in a fashion similar to its depth, with an apparently similar power.

Except for subsequent collapse and modification, discussed in Chapter 8, the final crater dimensions are established at the end of the excavation stage. Since some collapse occurs for all craters, the crater established at this stage is called the *transient crater* (Fig. 5.11d). The dimensions of the transient crater are conventionally given either as "apparent" depth and diameter,  $H_{at}$  and  $D_{at}$ , as earlier, which refer to the preimpact sur-



**Fig. 5.12** The stages of crater formation are reflected in the rate of change of depth with time. Crater depth increases linearly with time during the contact and compression stage. During the excavation stage crater depth increases as a power  $n$  of time, where  $n \approx 0.4$  in many experiments. Crater depth is constant after the end of the excavation stage, unless subsequent collapse decreases it during the modification stage. After Holsapple and Schmidt (1987).

face, or they are given as rim-to-rim diameter  $D_t$ , and the depth of the crater floor below the rim,  $H_t$ . The latter pair is more frequently used in geological work, since it can usually be measured more readily from photographs, whereas the apparent diameter and depths are easier to measure in the lab, where the preimpact surface is usually known.

The depth/diameter ratio of the transient crater,  $H_t/D_t$ , is typically between one-fourth and one-third for both experimentally produced craters and several intensively studied natural craters on earth (see Section 8.2). In the past, controversies have arisen over supposed changes (usually decreases) in the depth/diameter ratio of the transient crater as a function of increasing crater size, acceleration of gravity, or impact velocity. At the present time there is no unequivocal evidence that any of these supposed variations in depth/diameter are real; nor does theory give any support to such hypotheses. Large extraterrestrial craters that have been preserved from major collapse by low gravity, such as the 90-km diameter crater Pan on the small Jovian satellite Almathea, seem to have depth/diameter ratios comparable to those of small laboratory craters.

#### 5.5.4 Depth of excavation

Figure 5.13 illustrates the geometrical relations among the initial positions of material excavated from the target, material displaced downward into the crater floor, and material vaporized or melted by the shock wave. The final transient crater's walls and floor are formed by the expanded and displaced boundary between vapor, which escapes from the crater, and melt, which lines its interior throughout its growth.

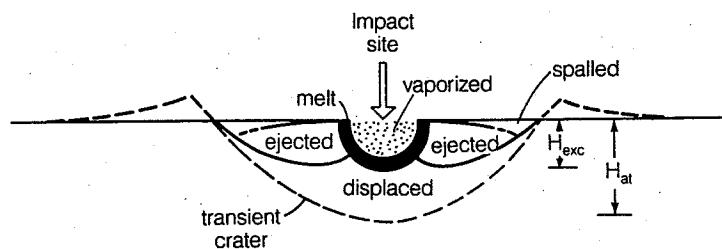
It is important to note from Figure 5.13 that the ejecta do *not* include material excavated from the full depth of the crater. This fact has also been demonstrated experimentally from work such as that illustrated in Figure 5.14, in which it is clear that the most deeply excavated horizon is only

one-third to one-half the depth of the transient crater below the preimpact surface. Target material deeper than the maximum depth of excavation is displaced downward beneath the crater floor: it does not emerge in the ejecta to be deposited on the target's surface. Strata below the maximum depth of excavation are thus pushed downward (note that in the experiment illustrated in Fig. 5.14 some of this displacement is accommodated by closing of intergranular pore spaces), whereas strata above this depth are bent upward. A transition zone develops between the deep and shallow regions around the crater. Above the transition zone, target material streams radially outward and upward along the wall of the growing crater, mixing turbulently with underlying, slower-moving debris. Melt ejected from this region is thoroughly mixed with less shocked debris. In the deeper region where material is simply displaced less mixing occurs.

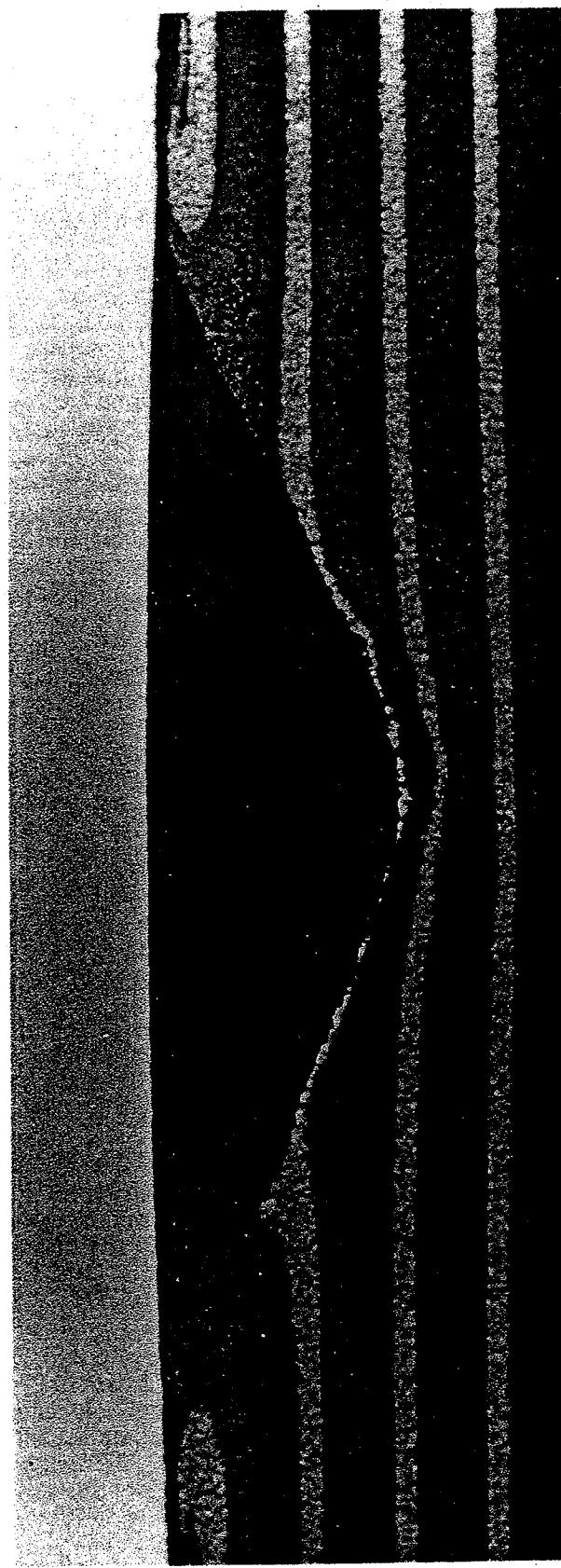
In the past there has been much confusion between the depth of excavation and the depth of the transient crater. Many lunar scientists were disappointed when the Apollo missions failed to return lunar mantle rocks from the ejecta blanket of the Imbrium basin on the moon. However, it is now clear from the geometry of the excavation flow that the maximum excavation depth is considerably shallower than the maximum depth of the crater. Although more study of the excavation flow may be necessary, a conservative rule of thumb based on computations and experiments is that the maximum depth of excavation  $H_{exc}$  is about one-third of the transient crater depth or one-tenth of the transient diameter:

$$H_{exc} \approx \frac{1}{3}H_t \approx \frac{1}{10}D_t \quad (5.5.4)$$

Note that the depth and diameter of large complex craters or basins must be corrected back to their transient values before the depth of excavation can be reliably estimated by this rule of thumb (see Section 8.3.3). Failure to make this correction will overestimate the excavation depth.



**Fig. 5.13** Provenance of material expelled from the crater. Vaporized material expands outward in the vapor plume. Of the remaining material some is ejected and some is displaced out of the crater and deforms the adjacent rocks, uplifting the surface near the rim and downwarping rocks beneath the crater. The ejected material is excavated from a maximum depth  $H_{exc}$  that is only about one-third of the transient crater depth or one-tenth of the transient crater diameter. The dashed lines show the profile of the transient crater.



**Fig. 5.14** Cross-section of a small-scale impact crater produced in a layered noncohesive sand target. The overturned flap near the rim, uplift of beds near the rim, and downwarp of beds beneath the crater are clearly seen. A white sand layer whose top is about one-third of the transient crater depth below the surface marks the transition between excavation and displacement. Photo courtesy of P. H. Schultz.

### 5.5.6 Maxwell's Z-model of the excavation flow

The details of the excavation flow can be determined only by experiment or by elaborate numerical computations. Even such numerical work may have difficulty in correctly computing the final dimensions of the transient crater. However, a simple analytical model of the excavation flow was proposed in 1973 by D. Maxwell and K. Seifert. This model gives a useful kinematic, if not dynamic, description of the cratering flow field, although, like all approximate models, it should not be used to determine fine details.

Maxwell and Seifert noted that in explosion cratering computations the radial component of the excavation flow velocity  $u_r$  usually falls as a simple inverse power of distance  $r$  from the explosive charge.

$$u_r = \alpha(t)/r^Z \quad (5.5.5)$$

where  $\alpha(t)$  is a function of time giving the strength of the flow and  $Z$  is the dimensionless power.

The incompressibility of the excavation flow,  $\nabla \cdot \mathbf{u} = 0$ , requires that the angular component of the flow velocity  $u_\theta$  in polar coordinates  $(r, \theta)$ , is

$$u_\theta = u_r (Z - 2) \sin \theta / (1 + \cos \theta) \quad (5.5.6)$$

The geometry of the velocity field defined by this model, seen in Figure 5.15, is remarkably similar to that computed in both explosion and impact cratering events for  $Z \approx 3$ . The equation of streamlines in polar coordinates is

$$r = r_0(1 - \cos \theta)^{1/(Z-2)} \quad (5.5.7)$$

where  $r_0$  is a constant that is different for each streamline. It is equal to the radius at which the streamline emerges from the surface ( $\theta = 90^\circ$ ). Taking  $r_0 = D_{\text{av}}/2$ , the radius of the transient cra-

ter, the maximum depth of excavation  $H_{\text{exc}}$  is

$$H_{\text{exc}} = (D_{\text{av}}/2)(Z - 2)(Z - 1)^{(1-Z)/(Z-2)} \quad (5.5.8)$$

For  $Z = 3$  the maximum depth of excavation  $H_{\text{exc}} = D_{\text{av}}/8$ , or about one-third of the final transient crater depth.

The total mass ejected from a crater described by the  $Z$  model,  $M_{\text{ej}}$ , is a fraction of the total mass displaced from the transient crater  $M_c$ .

$$M_{\text{ej}} = [(Z - 2)/(Z + 1)]M_c \quad (5.5.9)$$

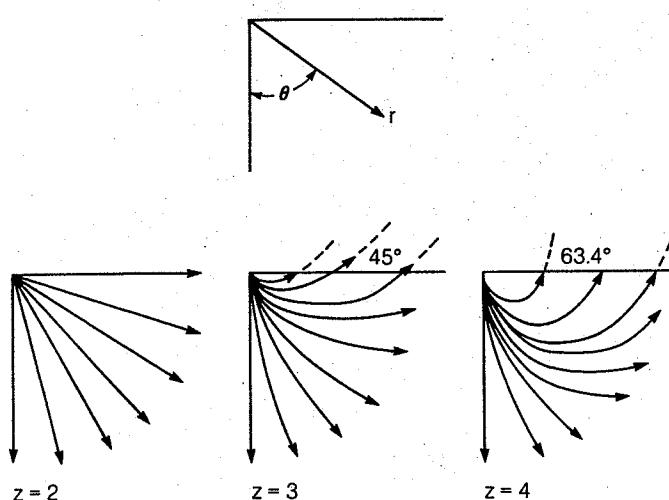
The  $Z$  model also predicts that the vertical and horizontal velocity components  $u_v$  and  $u_H$  of the ejecta launched at a distance  $s$  along the surface from the impact point are

$$u_v = \alpha/s^Z \quad (5.5.10)$$

$$u_H = (Z - 2)u_v \quad (5.5.11)$$

or that the angle of ejection is  $\phi = \tan^{-1}(Z - 2)$ , equal to 45 degrees for  $Z = 3$ .

The  $Z$  model presented thus far is a kinematical model useful for describing the form of the excavation flow. Maxwell and Seifert attempted to give it more dynamical content by computing the function  $\alpha(t)$  in Equation 5.5.5. This function gives the strength of the flow at any particular time. Its value is different for each streamtube in the flow. It is estimated by using energy conservation in each of these streamtubes, neglecting interactions with adjacent tubes. Thus, the sum of kinetic, gravitational, and distortional energy is found in each streamtube at some initial time. The total energy in each streamtube is conserved as the flow progresses. However, the kinetic energy declines at the expense of the gravitational and distortional energies, so that the net flow velocity declines.



**Fig. 5.15** Excavation flow fields for the  $Z$ -model. The streamlines for  $Z = 2$  (purely radial flow),  $Z = 3$ , and  $Z = 4$  are shown.  $Z = 3$  has been found to be a good approximation to actual impact cratering flows.

Once  $\alpha(t)$  is known for a given streamtube, the position of any point within the tube can be found by direct integration. The radial position  $r(t)$  of any point as a function of time is given by

$$r(t)^{Z+1} = r_0^{Z+1} + (Z+1) \int_0^t \alpha(t) dt \quad (5.5.12)$$

where  $r_0$  is the initial position of the point at  $t = 0$ . If  $r_0$  is the radius of the crater cavity's surface and if  $t = 0$  is the end of the contact and compression stage, then  $r(t)$  is the radius of the growing crater cavity throughout the excavation stage.

The depth/diameter ratio of the transient crater can be determined by integrating Equation 5.5.12 for the axial (vertical,  $\theta = 0$ ) streamtube, which yields the crater's depth,  $H_{at}$ , and dividing it by  $r(t)$  for the last streamtube in which the velocity magnitude  $\alpha(t)$  is positive, which yields the diameter  $D_{at}$ . This depth/diameter ratio for the "apparent" crater variables must be corrected for the rim height and slope to give the depth/diameter ratio  $H/D$ , in terms of geologic variables.

The Z-model described here can be (and has been) improved and extended in several ways. One of the most straightforward is to move the source of the flow,  $r = 0$ , from the surface to some depth below the surface, taking into account the depth  $d_{eq}$  of the effective center of the shock wave. A reference to work in which this is done (Croft, 1980) is given at the end of the chapter. Other workers have attempted to refine Maxwell and Siefert's methods of estimating energies in the streamtubes. The Z-model, however, is fundamentally limited by its neglect of interactions between the streamtubes. For this reason it can never become an exact description of the cratering flow, however accurately the dynamics within a single streamtube is represented.

The Z-model's accuracy has been tested against the results of several numerical computations. It was invented to describe the results of surface explosion cratering computations, and it works well for these with  $Z \approx 2.7$ . It also seems to represent impact cratering flow fields fairly well, provided the source of the flow is moved downward from the surface to an equivalent depth of burial. In detail, however, the best-fit value of  $Z$  is a function of both angle  $\theta$  and time. In one comparison (Austin et al., 1981),  $Z$  varied between about 2 near the axial streamtube,  $\theta = 0$ , to about 4 near  $\theta = 90$  degrees, with  $Z \approx 3$  as a reasonable overall approximation. The value of  $Z$  also increases with time according to this simulation. Furthermore, the Z-model fails at both very small and very large radii, compared with the transient crater radius.

In spite of all its faults, the Z-model gives a reasonably accurate representation of the gross features of the cratering flow and can even be used to predict some first-order dynamical properties,

such as the transition between strength-dominated and gravity-dominated craters or the dependence of crater diameter on gravitational acceleration (O'Keefe and Ahrens, 1981). It has the unfortunate feature of not being a truly dynamical model, so that further refinements are not necessarily closer approximations to the full dynamical equations of motion. Nevertheless, the excellent properties of this model are probably still far from being fully exploited.

## 5.6 COMPLICATIONS OF THE EXCAVATION STAGE

### 5.6.1 Oblique impact

The effect of oblique impact on the excavation of an impact crater is not well understood. Numerical studies cited in Section 4.3 show that the shock wave produced in an oblique impact is roughly hemispherical in spite of the obliquity, although it is weaker than the shock wave produced by a vertical impact at the same velocity. In confirmation of this observation, craters resulting from oblique high-velocity impacts are circular down to very low angles of incidence. Figure 5.16 shows that a series of small experimental craters formed in noncohesive quartz sand by projectiles traveling between 3.6 and 7.2 km/second were circular at all angles down to about 10 degrees. At smaller angles the craters became elongated along the projectile's flight direction. An example of an elongated crater on the moon is illustrated in Figure 6.9b.

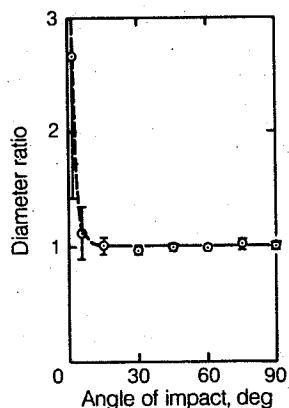


Fig. 5.16 Elongation of crater rim crests as a function of angle of impact. The vertical axis is the ratio between the downrange diameter and the crossrange diameter of the crater, and the horizontal axis is the angle of impact in degrees (90° is a vertical impact). The data is derived from 121 rounds of aluminum and pyrex spheres fired into noncohesive quartz sand at velocities ranging from 3.6 to 7.2 km/second. The craters are circular down to angles of incidence less than 10°. After Gault and Wedekind (1978).

Although the craters produced by oblique impacts are circular down to very low incidence angles, their ejecta may betray their obliquity. In the early part of the excavation stage, the projectile plows through the target, preferentially throwing material to the sides, perpendicular to its flight path. Very little material is thrown out behind the projectile, resulting in a characteristic wedge-shaped "forbidden zone" in the ejecta deposit, discussed further in Section 6.3.6 and illustrated in Figure 6.9a. At very low angles a downrange forbidden zone develops as well. These ejecta patterns indicate a corresponding asymmetry in the early part of the excavation flow, although in the case of a circular final crater the asymmetry evidently dies out before the rim forms.

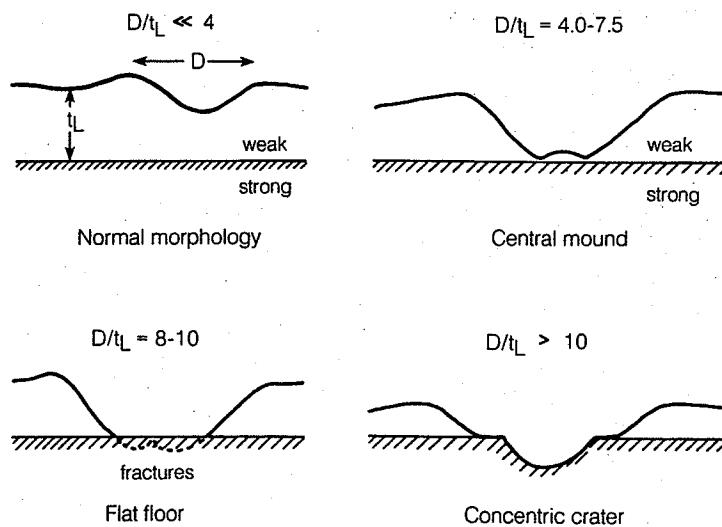
### 5.6.2 Layering and structure in the target

No geological target is composed of uniform, isotropic, and homogeneous material with a perfectly plane surface. Real targets may be layered or contain irregularly distributed rocks with different mechanical properties; are usually traversed by joints, cracks, and other planes of weakness; and often have significant topographic variations across the site of the eventual crater. The effect of these irregularities on crater formation has received only a small amount of attention, however.

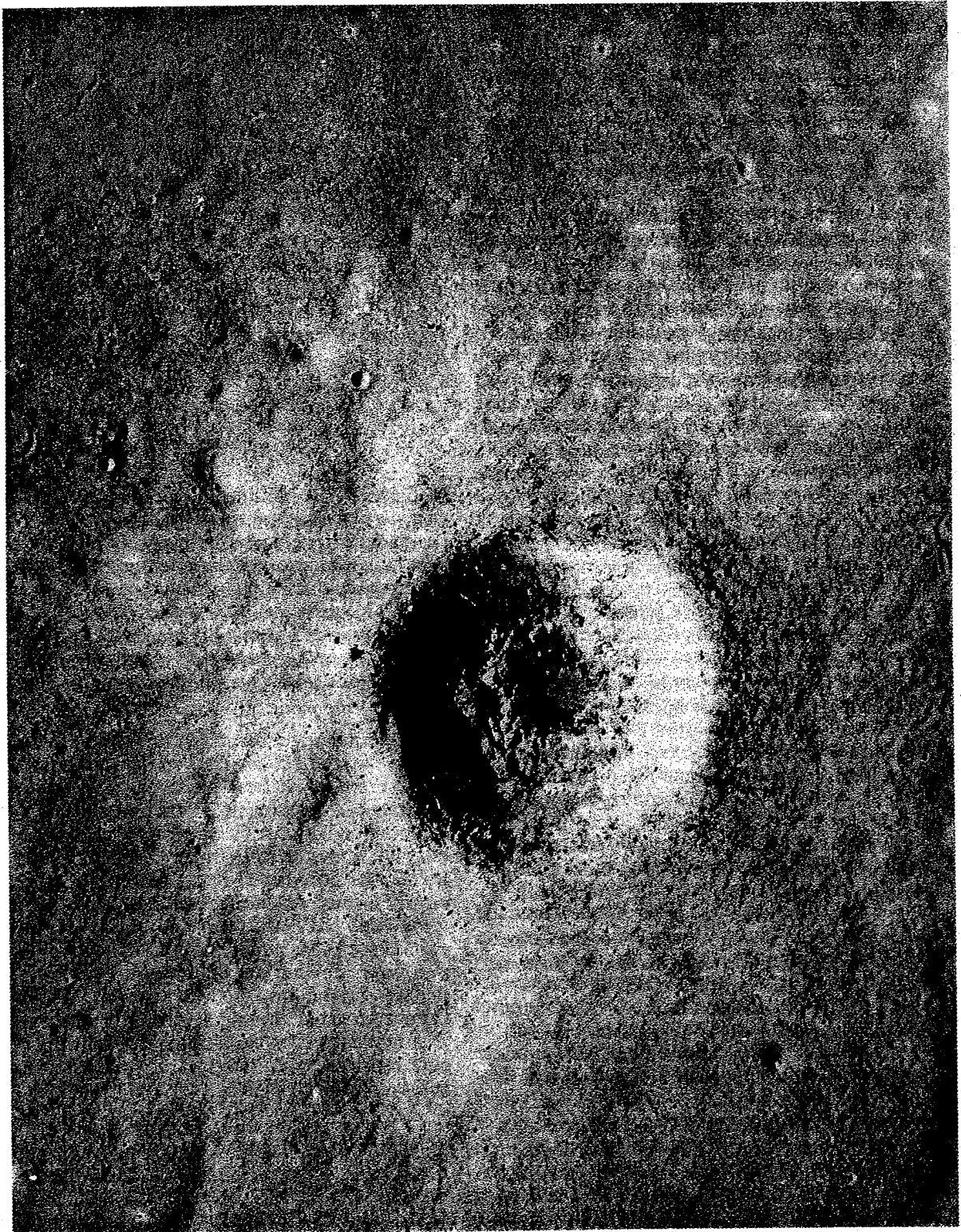
The most fully studied case is the effect of layering when a weak layer overlies a strong half space. This situation is common on the moon, where a loose, fragmental regolith overlies consolidated lava flows in the mare regions. This situation was studied experimentally by V. R. Oberbeck and W. L. Quaide in the late 1960s. They

found that the morphology of the final crater depends sensitively on the ratio between the crater's rim-to-rim diameter  $D$  and the layer thickness  $t_L$  (Fig. 5.17). When  $D/t_L < 4$  the crater has a normal bowl-shaped morphology. When  $4 < D/t_L < 7.5$ , however, a low central mound forms in the bottom of the crater on the strong layer. This mound is composed of unexcavated weak material. The streamtubes that would normally have carried this material farther down or out of the crater intersect the strong layer and motion in them stops sooner than for the rest of the flow. When  $8 < D/t_L < 10$ , the flow is strong enough to excavate the weak material from the crater floor, which is flat, constituting an exposure of the strong layer's top. Finally, for  $D/t_L > 10$  a crater forms in the strong layer as well, but the strength transition is demarcated by a bench or shelf on the wall of the final crater at the horizon of the strong layer. As the crater diameter increases beyond this final transition, the bench occurs progressively higher up on the crater wall and becomes less conspicuous. Figure 5.18 illustrates a wide bench low on the wall of a 1.2-km diameter lunar crater that is probably caused by a resistant rock layer about 100 m below the surface.

This set of morphologic changes that are due to layering was recognized in small lunar craters by Oberbeck and Quaide and was successfully used by them to estimate the regolith thickness on the moon before the Apollo landings. Small craters on Mars and even its satellite Phobos show similar morphologies and may thus be used to estimate regolith thickness. Benches formed on the walls of very large craters by layers of increasing strength have been suggested to explain the origin of multiring basins in the "nested crater model"



**Fig. 5.17** The morphology of craters formed in a weak layer overlying a stronger layer. Depending on the ratio between crater diameter  $D$  and weak layer thickness  $t_L$ , the crater may be a normal bowl shape, have a low central mound, flat floor, or interior benches on its walls. *After Quaide and Oberbeck (1968).*



**Fig. 5.18** This 1.2-km diameter crater on the moon shows a wide bench low down on its wall, suggesting that the moon's surface in this area consists of a weak layer about 100 m thick overlying a more resistant rock unit. Blocks up to 30 m in diameter litter the crater rim and a small smooth deposit in the very bottom of the crater may be impact melt. *Apollo panoramic photograph AS15-9287.*

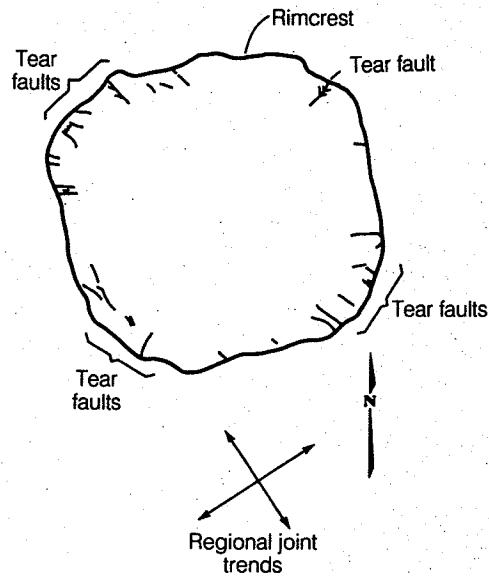
of ring formation, Section 9.4.3, although this is probably not how multiring basins form.

There has been little study of what happens when a strong layer overlies a weak one. Some small-scale impact experiments with a cohesive layer over loose sand show little change in overall morphology, although breakup of the strong surface layer results in larger clumps in the ejecta, blocks on the crater rim, and a more irregular outline of the crater rim in plan view. The presence of multiple layers in the target or of lateral strength variations have likewise received little study. One noteworthy observation from terrestrial explosion cratering experiments is that a water table in a noncohesive target acts very much like the strong layer beneath a weak layer described above. It produces the corresponding central peaks, flat floors, or wall benches, depending on the ratio between crater size and the depth to the water table. It is possible that morphologic variations in small Martian craters might be due to a water table or to a frozen layer in permafrost, but it seems impossible to differentiate the effect of such a water table from that of a resistant rock unit.

Joints, faults, or planes of weakness in the target rock also play a role in affecting the final crater form. Small-scale experiments (Gault et al., 1968) show that both the final crater's shape and the ejecta deposits are modified when such structures are present. Azimuthal variations of ejecta thickness are more pronounced when planes of weakness occur in the upper layers of the target, a circumstance that is signalized by the appearance of blocks on the crater rim. One of the most striking examples of structural control, however, is Meteor Crater, Arizona. Seen in plan view, the crater is actually more square than circular (Fig. 5.19). Small tear faults occur at each "corner" of the rim. These faults formed when adjacent plates of rock were uplifted different distances by the cratering flow. Meteor Crater's peculiar shape is caused by two orthogonal sets of vertical joints traversing the sedimentary rocks in which it formed. The joint directions form approximate "diagonals" across the crater. The cratering flow exploited these planes of weakness and peeled back the surface rocks much as the petals of a flower open, tearing along the joint directions.

Lines of weakness also influence the shape of much larger craters, but in a very different way, when collapse occurs. Rim collapse is much like a landslide, the headscarp of which parallels any lines of weakness in the source area. In these larger, collapsed (complex) craters (see Sec. 8.3), straight segments on the crater wall are thus parallel to lines of weakness, behaving very differently from small uncollapsed (simple) craters like Meteor Crater in which the straight segments bisect the joint trends.

Original topography in the target may also af-



**Fig. 5.19** The outline of Meteor Crater's rimcrest is more square than circular, betraying the presence of structure in the target rocks. The "corners" of the crater are cut by small tear faults across which differential uplift of the rim took place. Two prominent orthogonal joint sets occur in this area with trends, shown in the figure, cutting approximately diagonally across the crater's square shape. *Data from Shoemaker (1963).*

fect the final crater form, but this variable is even less well studied than layering and structure. Small craters that form on slopes are sometimes elongated in the downhill direction and may suffer more collapse and slumping on the uphill crater wall. A number of large lunar craters, such as the 75-km diameter King Crater (Fig. 2.10), clearly formed on surfaces with large topographic variations, but little change in overall circularity can be seen. The major effect observed in complex craters is that the slump terraces adjacent to the high topography are wider than any others in the crater. Since the original transient craters are greatly enlarged by slumping, however, it is difficult to deduce what effect, if any, the topography had on crater excavation.

#### FURTHER READING

##### *General References*

There is no comprehensive review of the excavation stage of impact cratering. The short description of the entire cratering process by Gault, Quaide, and Oberbeck cited at the end of Chapter 4 is the best general description. A rather specialized, but useful, review emphasizing computer models is

K. N. Kreyenhagen and S. H. Schuster, Review and comparison of hypervelocity impact and explosion cratering calculations, in

D. J. Roddy, R. O. Pepin and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 983-1002 (1977).

Another excellent review that emphasizes the relation between impact and explosion cratering is H. F. Cooper, A summary of explosion cratering phenomena relevant to meteor impact events, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 11-44 (1977).

A historically rather interesting analysis of impact crater excavation from an unconventional point of view deserves more attention than it has received:

E. J. Öpik, Meteor impact on solid surface, *Irish Astron. J.* 5: 14-33 (1958).

### Special Topics

No papers, to my knowledge, have focused exclusively on the structure of the shock wave produced by an impact. The structure of the shock wave is well delineated incidentally as part of a more general analysis of cratering flows by

M. G. Austin, J. M. Thomsen, S. F. Ruhl, D. L. Orphal, and P. H. Schultz, Calculational investigation of impact cratering dynamics: Material motions during the crater growth period, in *Proc. Lunar Planet. Sci. Conf. 11<sup>th</sup>*, pp. 2325-2345 (1980).

Some experimental data on shock-wave structure is presented in

J. Lipkin and M. E. Kipp, Wave structure measurement and analysis in hypervelocity impact experiments, *J. Appl. Phys.* 47: 1979-1986 (1976).

The relation between the shock wave and the excavation flow is the subject of

H. J. Melosh, Impact cratering mechanics: Relationship between the shock wave and excavation flow, *Icarus* 62: 339-343 (1985).

Information about ground motion near impacts can be derived indirectly from the much more abundant information on ground motion near explosions. One of the more readily available reviews of this subject is

H. F. Cooper and F. M. Sauer, Crater-related ground motions and implications for crater scaling, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 1133-1163 (1977).

Energy partition and the Gault-Heitowit approximation are the subject of the rather difficult-to-find 1963 report by Gault and Heitowit cited at the end of Chapter 4. A much more readily available reference is

S. W. Kieffer and C. H. Simonds, The role of volatiles and lithology in the impact cratering

process, *Rev. Geophys. Space Phys.* 18: 143-181 (1980).

The elastic energy radiated from an impact and the seismic shaking it produces are discussed in detail by

P. H. Schultz and D. E. Gault, Seismically induced modification of lunar surface features, *Proc. Lunar Sci. Conf. 6<sup>th</sup>*, pp. 2845-2862 (1975).

Vapor plume expansion from the site of an impact has not yet been much explored. One of the first important papers on the subject is

A. M. Vickery, Effect of an impact-generated gas cloud on the acceleration of solid ejecta, *J. Geophys. Res.* 91: 14,139-14,160 (1986).

The thermodynamics of vapor expansion into a vacuum is not usually treated in textbooks. An excellent review of the subject, including two-phase flow and condensation, is part of the rocket motor literature:

D. Altman and J. M. Carter, Expansion Processes, in D. Altman, J. M. Carter, S. S. Penner, and M. Summerfield (Eds.), *Liquid Propellant Rockets*, Princeton University Press, Princeton, N.J., pp. 26-63 (1960).

A detailed discussion of stress wave interference in the near-surface zone near an impact is presented in

H. J. Melosh, Impact ejection, spallation and the origin of meteorites, *Icarus* 59: 234-260 (1984).

The geometry of the excavation flow was part of an important paper that established the depth of excavation of impact cratering:

D. Stöffler, D. E. Gault, J. Wedekind, and G. Polkowski, Experimental hypervelocity impact into quartz sand: Distribution and shock metamorphism of ejecta, *J. Geophys. Res.* 80: 4062-4077 (1975).

A more recent discussion of the depth of excavation based on terrestrial crater data is

R. A. F. Grieve, P. B. Robertson, and M. R. Dence, Constraints on the formation of ring impact structures, based on terrestrial data, in P. H. Schultz and R. B. Merrill (Eds.), *Multi-ring Basins*, *Proc. Lunar Planet. Sci. 12A*, pp. 37-57 (1981).

The Maxwell Z-model was first described in a limited circulation Physics International Company report. A useful summary of the model was printed in

D. E. Maxwell, Simple Z model of cratering, ejection, and the overturned flap, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 1003-1008 (1977).

The implications of the Z-model for planetary cratering phenomena, including an appendix that presents many useful formulas for a Z-model flow centered some distance below the surface is found in

## IMPACT CRATERING

S. K. Croft, Cratering flow fields: Implications for the excavation and transient expansion stages of crater formation, *Proc. Lunar Planet. Sci. Conf. 11<sup>th</sup>*, pp. 2347-2378 (1980). The interesting effects of weak-over-strong layering in the target are discussed in the comprehen-

sive experimental and observational investigation of

W. L. Quaide and V. R. Oberbeck, Thickness determinations of the lunar surface layer from lunar impact craters, *J. Geophys. Res.* 73: 5247-5270 (1968).

## VI. EJECTA DEPOSITS

### 6.1 BALLISTIC FORMULAE

Most of the debris ejected from an impact crater falls back onto the surface of the planet on which it forms. Only a small fraction of the total mass ejected escapes the target planet's gravitational field. The material that falls back to the surface piles up in a series of deposits whose morphology depends upon distance from the impact site, crater size, substrate characteristics, and the presence or absence of an atmosphere or volatiles, particularly liquid water.

Most ejecta are emplaced ballistically by the free, or nearly free, flight of debris fragments expelled from the crater. The ballistic range-velocity formula is particularly important for interpreting ejecta deposits. A convenient form valid for short ballistic ranges  $R_b$  is

$$R_b = (v_e^2/g) \sin(2\Phi) \quad (6.1.1)$$

where  $v_e$  is the ejection velocity,  $g$  is the planet's surface acceleration of gravity, and  $\Phi$  is the ejection angle. The time of flight  $T_f$  is given by a similarly simple equation:

$$T_f = (2v_e/g) \sin \Phi \quad (6.1.2)$$

When the range becomes a significant fraction of the target planet's radius,  $R_p$ , the range equation becomes slightly more complicated:

$$R_b = 2R_p \tan^{-1} \left( \frac{(v_e^2/R_p g) \sin \Phi \cos \Phi}{1 - (v_e^2/R_p g) \cos^2 \Phi} \right) \quad (6.1.3)$$

This equation is equal to Equation 6.1.1 at short ranges,  $R_b \ll R_p$ . The equation for time of flight in this case is much more complex and because of its limited usefulness is not given here. The interested reader can find it in general works on ballistics or in Ahrens and O'Keefe (1978).

The exposition in this chapter proceeds from the crater rim outward to the most distant ray deposits. Most examples are drawn from the moon or the terrestrial planets because they have been imaged at the highest resolution. Ejecta on the icy satellites of Jupiter are discussed briefly. Much remains to be learned about ejecta deposits on small bodies such as the Saturnian satellites or the asteroids, although the physics of emplacement should be the same as on any other body.

### 6.2 CRATER RIMS

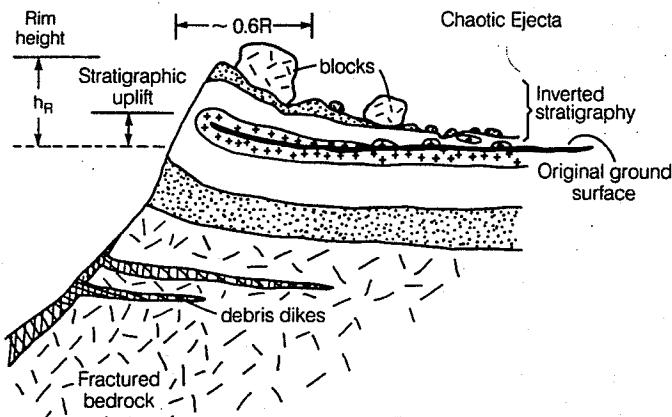
The earliest telescopic observations of lunar craters revealed that the crater cavity is encircled by a rim that stands above the surrounding terrain. Little progress was made in understanding the nature of these rims, however, until the early 1960s when geologic studies of Meteor Crater, Arizona, and nuclear explosion craters were performed, principally by E. M. Shoemaker.

These studies showed that the crater rim is not entirely composed of ejected debris. Approximately half the height of the rim crest is due to structural uplift of the underlying preimpact surface. This uplift is largest beneath the rim crest and dies off rapidly with increasing distance from the crater. It drops to zero approximately 1.3 to 1.7 crater radii (center-to-rim-crest radius) from the crater's center. This uplift is mainly due to plastic deformation of the rock below the rim. Strong horizontal compressive forces push outward from the crater's center as it grows, fracturing the rock and squeezing it upward. These forces may also produce small inward- or outward-dipping thrust faults, as at the 167-m diameter Odessa, Texas, impact crater or at Meteor Crater. Concentric rim anticlines around both the Odessa Crater and Jangle U nuclear explosion crater were pushed up by the same forces.

Another contributor to structural rim uplift is injection of subhorizontal breccia dikes into the crater wall. At Meteor Crater these dikes carry brecciated rock and meteorite fragments well beneath the crater rim. Laboratory and field investigations of explosion craters show that the rim of the transient crater is initially thrown several times higher than its final position, then settles back under gravity. During this time the vertical stress beneath the rim is less than hydrostatic. The horizontal stress, however, is still compressive. This stress field injects debris in the crater cavity into subhorizontal dikes beneath the rim.

The remainder of the rim's height is built up from ejected debris that lies on top of the uplifted ground surface. This material is thrown out at such low velocity that it frequently retains its stratigraphy and, where good stratigraphic control exists, may be recognized as an "overturned flap" of inverted rock units. Figure 6.1 illustrates the overturned flap and structural uplift in the rim of a

## IMPACT CRATERING



**Fig 6.1** Schematic cross section of the rim of a small impact crater. This figure illustrates the stratigraphic uplift, confined to about  $1.6 R$  of the crater's center, where  $R$  is the crater radius. This uplift is caused by dikes injected beneath the crater rim and plastic deformation of the underlying rocks. The remainder of the rim height is made up of ejecta deposited on the original ground surface. Near the rim the original stratigraphy of the surface rock units is inverted. Large blocks lie on the rim deposits.

hypothetical small impact crater. The "rim syncline" or overturned fold shown in the figure is not always present because collapse frequently carries it downward into the crater cavity, even in small craters. When this occurs, only an inverted stratigraphy that begins at the preimpact ground surface can be recognized.

Shoemaker's recognition of the overturned flap in the early 1960s became so famous that many geologists assumed that the ejecta blanket was little more than an upside-down stratigraphic section through the entire thickness of material intersected by the transient crater. The deficiency of this view was amply illustrated by the disappointment of many lunar scientists when samples returned from the Imbrium ejecta blanket proved not to be lunar mantle material. In fact, as explained in Section 5.5.4, only material shallower than roughly one third of the transient crater depth is ever ejected. The rest of the crater is excavated by displacement of target material downward and outward beneath the crater rim. Moreover, the overturned flap, however characteristic it may be of the crater rim, is not recognizable far beyond it. Ejecta landing at larger distances from the crater rim strike at high velocity and mix into a stratigraphic jumble that includes material scoured from the preexisting ground surface, as apparent in the ejecta of the 22-km diameter Ries Crater, Germany.

The total rim heights  $h_R$  of many lunar, terrestrial, explosion, and laboratory impact craters have been measured. These measurements are all consistent, within a factor of about 1.5, with a formula derived from lunar craters smaller than 15 km in diameter by Pike (1977):

$$h_R = 0.036D^{1.014} \quad (6.2.1)$$

where  $D$  is the crater rim-to-rim diameter and all units are in meters. The exponent of this equation

is not significantly different from 1, suggesting that all simple, uncollapsed craters have the same proportions, with a rim height equal to about 4 percent of the crater diameter. Craters larger than about 15 km in diameter on the moon show signs of collapse. Their rim heights are expressed by a very different relation, also from Pike (1977):

$$h_R = 0.236D^{0.399} \quad (6.2.2)$$

for  $D > 15$  km on moon

These rim heights are diminished over those predicted by Equation 6.2.1 because the transient crater's rim crest slides into the crater during collapse. This collapse may also carry both the overturned flap and the structurally uplifted rim into the interior of these craters. The coefficient in Equation 6.2.2, and perhaps the power of  $D$ , should vary from planet to planet because the degree of crater collapse depends upon the surface gravity and material (see Chap. 8). Equation 6.2.1, however, is valid for all uncollapsed craters, independent of planet.

The rims of craters formed in hard rocks are littered with blocks of rock. The blocks are usually largest on the rim because the ejection velocity is lowest there. Although the block size should be affected by variables such as rock strength, joint spacing, and perhaps impact velocity, a study of blocks on the rims of lunar craters ranging from a few meters to nearly 100 km in diameter yields a crude relation between maximum block size  $l_b$  and crater diameter for blocks larger than about 1 m (dimensions are in meters: Moore, 1971):

$$l_b \approx (0.1 \text{ to } 0.3)D^{2/3} \quad (6.2.3)$$

Observed block sizes often differ from this relation by a factor of two, perhaps reflecting the expected variations of rock strength or other impact conditions.

The rim blocks are emplaced by the low-velocity tail of the excavation flow. These blocks ascend along the crater's inner wall from some distance below the original surface, in consonance with the inverted stratigraphy of the rim region. These shattered fragments of the original subsurface rock layers are further reduced in size as they participate in the strongly sheared excavation flow and are thrown to their final resting places on the rim.

### 6.3 EJECTA BLANKETS

Essentially all impact craters are surrounded by a deposit of debris ejected from the crater interior.

The only exceptions are craters on steeply sloping surfaces or on satellites with too little gravity to retain the ejecta. This ejecta deposit is thickest at the crater rim and thins with increasing distance from the rim. Where this deposit is recognizably continuous near the crater it is called an *ejecta blanket*. Ejecta beyond the edge of the continuous deposit are thin and patchy. Secondary craters occur in this area and beyond it. Figure 6.2 shows the ejecta blanket of the 30-km diameter lunar crater Timocharis. The preexisting terrain within about one crater radius of the rim is buried and mostly obliterated by the continuous ejecta blanket. Light patches show where thinner deposits overlie the mare surface.

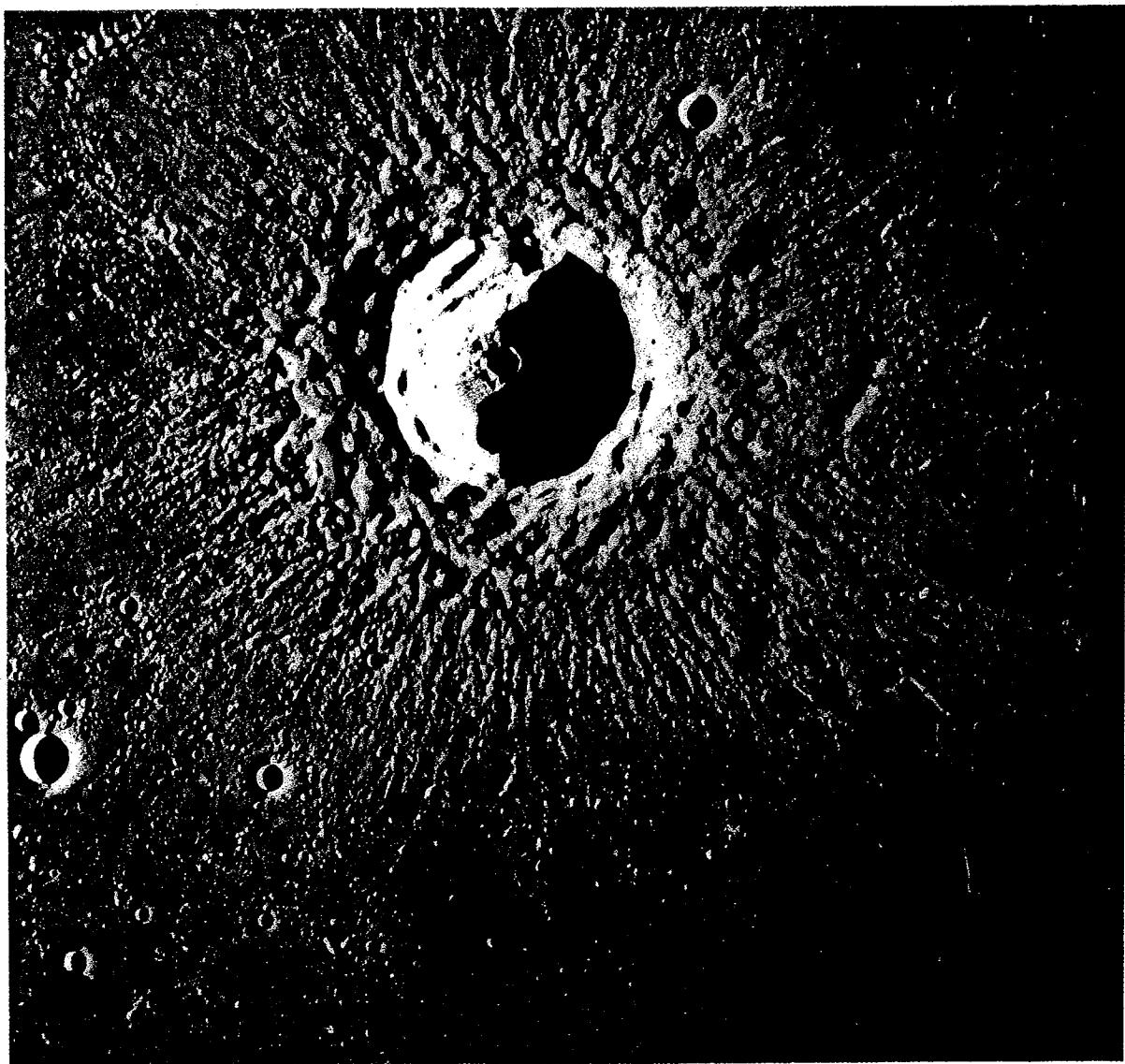


Fig. 6.2 Near-vertical view of the 30-km diameter lunar crater Timocharis. The ejecta deposits are dunelike near the crater rim but grade into a subradial facies beyond about  $2 R$  from the crater's center. Secondary craters occur at greater distances. A patch of bright material surrounding the crater indicates the presence of ejecta too thin to greatly modify the preexisting terrain. *Apollo Photo AS15-1005.*

### 6.3.1 Extent and thickness

Continuous ejecta extend about one crater radius from the crater rim, apparently regardless of crater size. A compilation of data on lunar craters by Moore et al. (1974) shows that the average radius of continuous ejecta  $R_{ce}$  surrounding craters of radius  $R$  is

$$R_{ce} = (2.3 \pm 0.5)R^{1.006} \quad (6.3.1)$$

This relation is valid for lunar craters ranging from about 1.3 to 436 km in diameter. It also seems to hold for much smaller craters in loose or unconsolidated targets. Ejecta blankets of small (strength-dominated) craters in rock do not follow this relation.

Ejecta blankets around craters on Mercury, Mars, and the Jovian satellites Ganymede and Callisto also obey the rule of thumb that the continuous deposits extend about one or two crater radii beyond the crater rim. Even the small Saturnian satellites Dione and Rhea have craters with recognizable continuous ejecta blankets that are geometrically similar to those of lunar craters. There is some suggestion (see, especially, Gault et al., 1975) that Mercurian ejecta blankets extend only 0.65 as far as lunar ejecta blankets, but this may be due to greater widening of the transient crater by slumping on this higher-gravity planet. Detailed studies of ejecta blanket sizes have been carried out only on the moon and Mercury, so it is difficult to make exact generalizations.

The study of the thickness and volume of ejecta blankets has been fraught with controversy. In work published between 1791 and 1802 the astronomer Schröter developed his "rule" equating the volume of the rim deposits to the volume of the crater bowl. Schröter's rule has turned out to be unverifiable because the deposits thin so gradually with increasing distance from the crater that significant volumes of material may be strewn discontinuously over rough terrain, making accurate volume estimates impossible. Only in recent explosion cratering experiments where the original ground surface is clearly marked can a form of Schröter's rule be verified. Such volume estimates must take account of rim uplift, expansion of the broken rock underlying the crater, and bulking of the loose ejecta. Other work on "ballistic sedimentation" (Sec. 6.3.3), moreover, indicates that the ejecta blankets of large craters may incorporate large amounts of substrate material, making the ejecta deposit volume significantly larger than the crater volume.

A widely cited compilation of data mainly from explosion experiments (McGetchin et al., 1973) states that the ejecta blanket thickness  $\delta$  is a power function of distance  $r$  from the crater center:

$$\delta = 0.14R^{0.74}(r/R)^{-3.0} \text{ for } r \geq R \quad (6.3.2)$$

where  $R$  is the transient crater radius and all dimensions are in meters.

Numerical calculations for shallow nuclear explosions, nuclear surface bursts, and impacts indicate some departure from the power -3 dependence of  $\delta$  on  $r/R$ , although experimental studies of small, half-buried explosive charges show that -3 works well. The power varies from -2.8 to -3.5 in small-scale impact experiments. A power of  $-3 \pm 0.5$  thus gives a good first estimate.

The coefficient of the power-law term in Equation 6.3.2 is more controversial. Thickness estimates of lunar basin ejecta made on the basis of this equation differ by a factor of five from geologic estimates in some places. Part of the reason for this disagreement is uncertainty in the transient rim diameter of large craters and, especially, of basins where confusion engendered by the multiple rings makes even the collapsed crater diameter uncertain by factors of two or more. Other important effects are the sometimes marked azimuthal variation of blanket thickness and the incorporation of substrate material into the blanket.

The best description of ejecta blanket thickness at the present time is the less definite formula

$$\delta = f(R)(r/R)^{-3.0} \text{ for } r \geq R \quad (6.3.3)$$

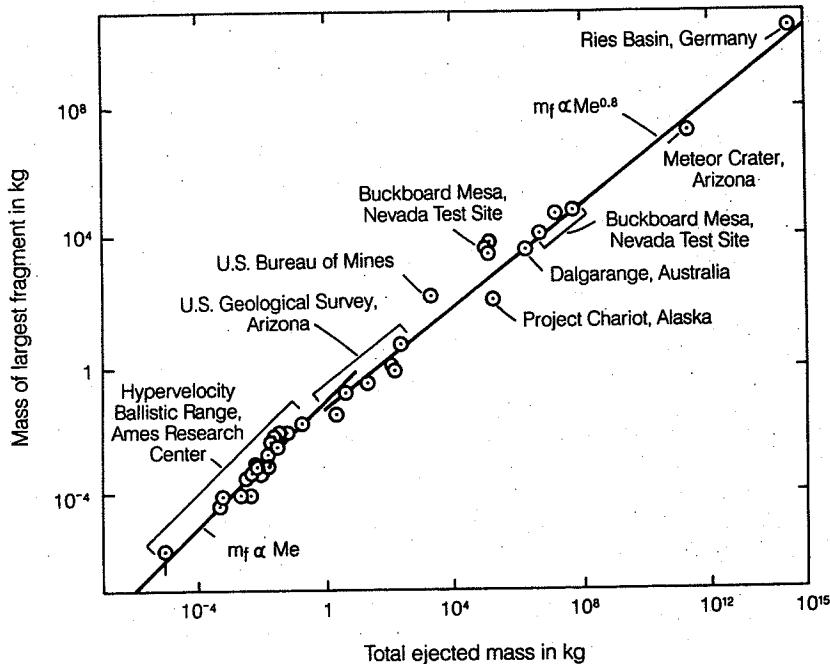
where  $f(R)$  is a poorly known function that is loosely constrained by the requirement that the total ejecta volume is at least as large as the volume of the crater below the preimpact ground surface.

Integration of Equation 6.3.3 shows that most of the ejecta lie close to the crater rim. The total volume of ejecta within a radius  $r$  of the crater's center falls as  $1/r$ , so that 50 percent of the ejecta lies between the crater rim at radius  $R$  and  $2R$ , 17 percent lies between  $2R$  and  $3R$ , and only 8 percent lies between  $3R$  and  $4R$ . This concentration of ejecta near the rim accounts for the near-complete blanketing of preexisting terrain and structures close to the crater that defines the "continuous" ejecta blanket.

### 6.3.2 Clast size distributions

Crater ejecta deposits are composed of vast numbers of broken rock fragments (clasts) mixed with a variable quantity of glass. Clast sizes range from microns up to many meters in diameter. Unlike many sedimentary rocks, no one size dominates the distribution, although small fragments greatly outnumber large ones.

There are several equivalent representations currently in use that describe how the number of fragments varies with fragment size. The most general and easily interpreted of these is the cumulative number  $N(m)$  of fragments of mass equal to or greater than mass  $m$ . Measurements on laboratory-scale impact craters, chemical and



**Fig. 6.3** The mass of the largest intact block in the ejecta blanket of a crater is a function of the total mass ejected for a variety of small-scale impact craters, explosion craters, and terrestrial impact structures. After Gault *et al.* (1963).

nuclear explosion craters, and the ejecta from large terrestrial and lunar impact craters are well expressed by a power-law relation

$$N(m) = C_f m^{-b} \quad (6.3.4)$$

where  $b$  commonly ranges between 0.8 and 0.9 and  $C_f$  is a constant. The size of  $b$  is thought to depend upon the fragmentation process: single crushing yields  $b$  values between 0.4 and 0.6, whereas  $b$  near 1 indicates multiple fragmentation.

Values of  $b$  in excess of 1 are occasionally observed in the deposits of impact craters. When  $b$  exceeds 1, the total mass  $M_T = \int m dN$  diverges at the low mass limit. This means that the power law (Eq. 6.3.4) cannot continue to arbitrarily small masses, but that it is cut off at some small mass below which the number of small fragments decreases rapidly with particle size. Measurements on nuclear explosion craters indicate that this cutoff takes place for particles in the 10<sup>-4</sup>-m to 10<sup>-5</sup>-m diameter range (10<sup>-9</sup> to 10<sup>-12</sup> kg).

The number distribution (Eq. 6.3.4) is also cut off at the high mass end. Measurements on numerous craters, illustrated in Figure 6.3, show that the mass of the largest ejected fragment  $m_f$  is a function of the total mass ejected  $M_e$ . For craters larger than roughly a meter in diameter this relation is

$$m_f = 0.8 M_e^{0.8} \quad (6.3.5)$$

By definition, the cumulative number of fragments  $N(m) = 1$  when  $m = m_f$ .

The number distribution, Equation 6.3.4, and maximum fragment mass, Equation 6.3.5, apply to the ejecta blanket taken as a whole. It is expected that fragment size decreases with increasing distance from the crater, an expectation that has been quantitatively verified in numerous small-scale impact experiments, at Meteor Crater, Arizona, and in the Bunte Breccia ejecta deposit of the 22-km diameter Ries Crater in Germany. The mean grain size of the Bunte Breccia falls as the inverse ninth power of the distance from the crater's center (Hörz *et al.*, 1983).

Distal ejecta are also expected to be, on average, more highly shocked than ejecta near the crater rim. However, because ejecta flow lines cut across pressure contours (See Sec. 5.5.2), ejecta at any given range are a mixture of materials shocked at different pressures. More distal ejecta merely contain a higher proportion of highly shocked or melted material than less distal ejecta.

Impact "melts" occurring in the ejecta of large craters are not uniform sheets of glass. All "melt sheets" that have been examined to date are chaotic mixtures of a chemically rather uniform melt glass with variable amounts of shocked fragmented ejecta. When the intimate mixture of hot melt and cold clasts strikes the surface it cools quickly, quenched by the clasts. The resulting material is probably incapable of flowing very far after it is emplaced, although in large craters where the pro-

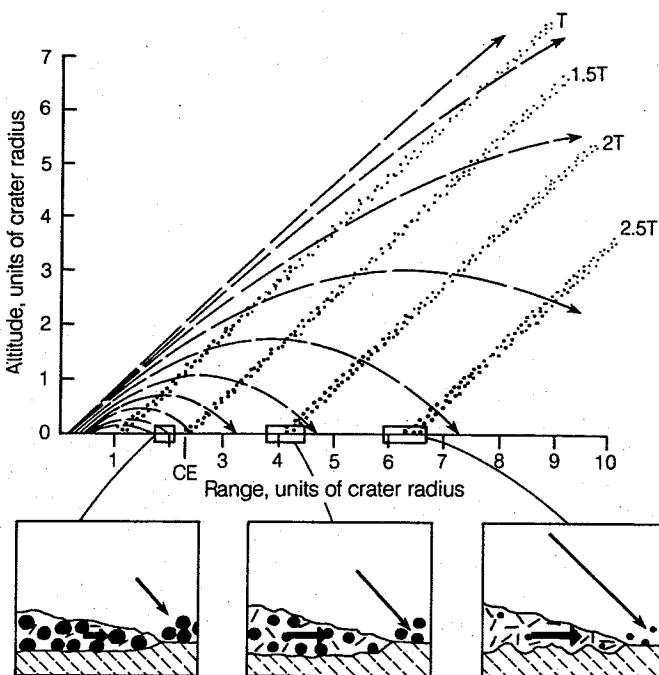
portion of melt to clasts is high, the clasts may be digested and flows then occur. In the Ries Crater the melt rocks (Suevite) overlie the Bunte Breccia deposits and evidently fell onto the main ejecta deposit after it had been emplaced. Deposition of the Suevite may, in this terrestrial crater, have been by a "base surge" type mechanism in which an air suspension of the droplets of hot melt and cold clasts played a major role.

### 6.3.3 Ballistic sedimentation

The ejecta deposited around an impact crater on an airless planet are emplaced ballistically; that is, ejecta are thrown from the crater with some initial velocity, follow a nearly parabolic trajectory above the planet's surface, then fall back to the surface, striking with the same velocity possessed upon ejection. Some interaction may occur be-

tween ejecta fragments in the denser parts of the ejecta curtain, but the general motion is dominated by ballistics alone. The ground surface around the crater is profoundly affected by the ejecta, and its interaction with this falling debris determines the character of the ejecta deposits of large craters.

The debris ejected from an impact crater travels together in the form of an "ejecta curtain." Although each fragment follows a parabolic trajectory, the times and velocities of ejection from the crater are organized so that most of the debris lies on the surface of an expanding inverted cone. Figure 6.4 illustrates the parabolic trajectories of a number of fragments from an impact crater. The positions of these fragments at several times are indicated. At any one time the debris lies on the surface of a cone that makes an angle of about 45 degrees with the ground surface. This cone sweeps



**Fig. 6.4** The debris ejected from an impact crater follows ballistic trajectories from its launch position within the final crater (the rim of the final crater is located at range equal to one crater radius). The innermost ejecta are launched first and travel fastest, following the steepest trajectories shown in the figure. Ejecta originating farther from the center are launched later and move more slowly, falling nearer the crater rim. Because of the relationship between the position, time, and velocity of ejection, the debris forms an inverted cone that sweeps outward across the target. This debris curtain is shown at four separate times during its flight, at 1, 1.5, 2, and 2.5  $T_f$ , where  $T_f$  is the crater formation time,  $(D/g)^{1/2}$ . Coarser, less shocked debris travels near the base of the curtain, whereas the fast, highly shocked ejecta fraction tends to travel near the top. The three insets show details of the preexisting ground surface when the ejecta curtain arrives. As range increases the ejecta strike with progressively larger velocity, incorporating larger amounts of the surface material and imparting a larger net horizontal velocity.

The ejecta curtain positions were computed from the theory of ejecta scaling (Sec. 7.12), assuming an ejection time  $t_{ej} = (r/R)^{2.8} T_f$ , ejection velocity  $v_e = 0.28(r/R)^{-1.8} \sqrt{gD}$ , and ejection angle of 45°. The tick mark labeled "CE" indicates the edge of the continuous ejecta. *Inspired by a figure in Oberbeck (1975).*

rapidly outward from the crater rim. Debris from the curtain strikes the ground at its base, impacting first near the crater rim, then at greater distances as time progresses. The size of ejecta fragments near the base of the ejecta curtain is expected to be larger than the fragments higher in the curtain, and the proportion of highly shocked fragments and glass increases with height in the curtain.

Terrestrial nuclear explosion craters and some volcanic craters of explosive origin are surrounded by ejecta deposits that superficially resemble the deposits of much larger extraterrestrial impact craters. Although there is a ballistic component to the emplacement of these deposits, much of the material is laid down by a *base surge*. A base surge is a gravity-driven density current that flows down from the rising mushroom cloud of a nuclear or volcanic explosion. This current is a mixture of air and dust originally entrained by the fireball that flows over the surrounding terrain in much the same way that muddy submarine turbidity currents flow out over the deep sea floor. As the base surge spreads out it slows, depositing its sediment load and producing high-velocity flow bedforms.

Base surges have been suggested as the mechanism by which impact crater ejecta deposits are emplaced. The heyday of the base surge idea was between about 1965 and 1975 when many lunar geologists found the idea appealing even though the moon has no atmosphere and density currents in a vacuum are hard to imagine. Base surges were popular because they explained many of the dune like forms and flow features seen in the ejecta deposits of large lunar craters and basins. No other mechanism for producing these features was known at the time.

A breakthrough in lunar sedimentology occurred in 1975 when V. R. Oberbeck studied the effect that the falling ejecta would have on the lunar surface. Oberbeck found that many of the flowlike features characteristic of ejecta blankets could be explained as resulting from the ejecta-surface interaction. He coined the term *ballistic sedimentation* to describe the process.

Ballistically emplaced debris falling near the crater rim strikes at low velocity because it travels only a short distance. At the rim itself this velocity is so low that rock units may retain some coherence and produce the overturned flap illustrated in Figure 6.1. At greater distances from the crater rim, the debris strikes at higher velocity. When this velocity is large enough, surface material is eroded and mixes with the debris. The falling ejecta also possess a radially outward velocity component. Although the vertical velocity is canceled when the debris strikes the surface, the outward momentum is retained by the mixture of debris and surface material (see the lower insets in Figure 6.4). This mixture moves rapidly outward

as a ground-hugging flow of rock debris, similar in many ways to the flow of a large dry-rock avalanche. Depositional features such as dunes, ridges, and radial troughs indicative of high-speed flow may result from this motion. The deposit itself consists of an intimate mixture of primary crater ejecta and of secondary material scoured from the preexisting ground surface.

The processes of ballistic erosion and sedimentation are usually unimportant for small craters because the impact velocity within two or three crater radii, where most ejecta lands, is too low for significant modification of the surface. For example, ejecta landing at an angle of 45 degrees two crater radii from the rim of a 100-m-diameter crater on the moon strikes at only 13 m/second. However, ejecta landing two crater radii from a 100-km diameter crater on the moon strikes at 410 m/second and may thus severely modify the surface. The horizontal component of this velocity is nearly 300 m/second, several times faster than the fastest-known terrestrial rock avalanche. (For example, the 1968 Peruvian Huascaran avalanche traveled at an estimated 78 m/second.) Although this velocity is reduced as momentum is transferred to surface materials incorporated into the flow, it is not surprising that the outer parts of the continuous ejecta blanket around large craters show features indicative of high-speed radial flow. Base-surge-type density currents are thus not necessary to explain these features.

Recent studies of the ejecta blanket of the Ries Crater in Germany (Hörz et al., 1983) strongly support the dominance of ballistic emplacement, even though the Ries almost certainly had a base surge associated with it. The evidence for a base surge is that the highly shocked and melted Suevite deposits overlying the main mass of the ejecta blanket have features suggestive of deposition from a base-surge-type flow. However, only 34 percent of the voluminous Bunte Breccia ejecta deposits are primary, clearly showing the dominance of ballistic sedimentation over base surge deposition.

Because a greater proportion of surface material is incorporated in the ejecta of larger craters by ballistic sedimentation, the ejecta thickness formula (Eq. 6.3.2) established for small craters probably underestimates the thickness of the ejecta blankets of large craters. The formula may, however, give a good estimate of the amount of primary ejecta at a given distance from large craters. The proportion of primary ejecta to material from the preexisting surface in ejecta blankets is presently not well known, except for the Ries Crater, although Apollo orbital geochemical investigations have shown that lunar basin ejecta often possess the same geochemical signatures as the underlying rocks, supporting the idea of large-scale incorporation of local material into the ejecta blankets of large craters and basins.

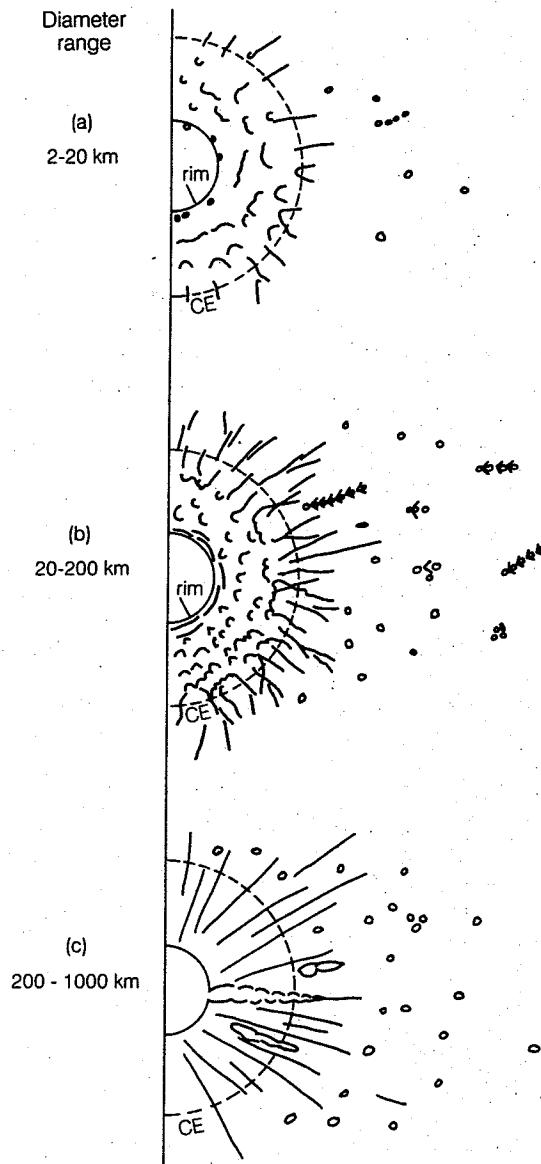
### 6.3.4 Ejecta facies

The topography of fresh ejecta blankets varies systematically both with distance from the crater and with primary crater size. These variations are mainly due to differences in ejecta velocity as a function of distance from the crater. A number of distinct facies, or topographic styles, were first described around fresh lunar craters (Howard, 1974). Similar facies occur around Mercurian and Martian craters, although additional types found on Mars are probably due to the presence of subsurface water excavated with the ejecta or to atmospheric interference with the ejecta curtain.

The ejecta deposits of the smallest class of lunar craters that penetrate the fragmental regolith, from 0.5 to about 1 km in diameter, are blocky and show little structure. Craters near the upper size limit of this class sometimes have subdued concentric dunelike structures close to their rims. Ejecta deposits of this type also occur around nuclear and chemical explosion craters. The ejecta of these small craters land at such low velocity that the preimpact surface is not significantly scoured and the ejected debris remains nearly where it falls.

Craters in the 2- to 20-km diameter range (Fig. 6.5a) have well developed concentric dunes beyond their rims (see Figure 2.2). The rims appear generally smooth to distances of  $1.2 R$  from the center, although they are often blocky when seen at high resolution. Rims may be locally veneered with thin melt-rock deposits. The dunes grade into a subradially braided fringe near the edge of the continuous ejecta deposit. These dunes superficially resemble barchan sand dunes whose crescentic "wings" face away from the crater's center. There is, however, no evidence that they were deposited from a moving fluid. Their arcuate forms have suggested to some lunar scientists that they are remnants of large secondary craters whose downrange portions were scoured away by ground-hugging ejecta flows. Alternatively, they could be produced by the unsteady, surging flow of the debris blanket as it moves outward under the impetus of individual clumps of falling ejecta. Debris is sometimes piled up against the craterward side of preexisting topographic obstacles, demonstrating that it flowed along the ground surface. The braiding near the outer margin of the ejecta blanket is probably due to discontinuous scouring of the surface by individual fragments or clumps of higher velocity ejecta. This radial facies is often lacking around the smaller craters in this diameter range ( $\leq 10$  km).

Hummocky terrain, consisting of concentrically aligned ridges and furrows, develops near the rims of craters between 20 and several hundred km in diameter (Fig. 6.5b). This terrain extends between 1.2 to 2.0  $R$  from the crater center. The rims are invariably slumped and at high



**Fig 6.5** Schematic illustration of ejecta facies for (a) craters in the 2- to 20-km diameter range, whose ejecta blankets are dominated by concentric dunes, (b) for craters in the 20- to 200-km diameter range, whose ejecta are characterized by an inner hummocky zone and outer radially lineated zone near the edge of the continuous ejecta, and (c) for craters larger than 200 km in diameter, which are surrounded by radially lineated ejecta in which several prominent troughs may occur. The larger craters are surrounded by secondary craters in clusters and chains.

resolution display fine concentric lineations that may be fractures related to the slump blocks. Pools of impact melt, flow fronts, and short leveed flow channels commonly occur on the rims and may be traced outward to distances of a crater radius. Concentric dunes similar to those in

smaller craters develop beyond the hummocks. The dunes grade into a zone of radial ridges that appear to be depositional.

These ridges may be similar in origin to the longitudinal striations on the deposits of large rock avalanches, for example, the 1964 Sherman Glacier landslide in Alaska (Shreve, 1966), or the deposit near Tsilokovsky on the moon (Howard, 1973). The exact cause of these avalanche striations is not known. They frequently occur on thick masses of rock debris that have flowed at speeds between 30 and 100 m/second without lateral confinement, conditions rather similar to those under which the radial ridges around craters develop.

Ejecta in the zone of radial ridges becomes discontinuous and grades into loops and clusters of secondary impact craters and herringbone structures (Sec. 6.4) that extend many crater radii from the impact. In the gradational region both ridges and secondaries modify each other's deposits. These relations may be clearly seen in the ejecta blankets of the 30-km diameter lunar crater Timocharis in Figure 6.2 and the 90-km diameter Copernicus in Figure 6.10.

The ejecta deposits of the largest, basin-forming, lunar impact craters are dominated by radial structures (Fig. 6.5c). Numerous radial grooves and gouges were first recognized around the Imbrium Basin by G. K. Gilbert in 1893, who named the topographic modification "Imbrium Sculpture." Imbrium sculpture extends more than 1000 km from the prominent Imbrium basin ring. Similar radial structures occur around the margins of all large craters, including the 1300-km diameter Caloris Basin (Fig. 9.6) on Mercury. Although these radial structures were at one time believed to be fractures, current opinion has returned to Gilbert's view that they are produced by ejecta from the primary impact event. Within about one crater radius of the rim the radial structures are depositional, and are evidently similar to the radial ejecta facies around smaller craters. Farther from the crater or basin center the grooves and gouges are created by coalescing chains of secondary craters. They range in size from valleys tens of kilometers across down to lineations only a few tens of meters wide.

Nearly every large crater or basin seems to have one or more especially prominent radial troughs extending from the rim out to nearly one crater diameter. The deep cleft known as the Alpine Valley on the north flank of the Imbrium Basin is an excellent example, but the ejecta blankets of Orientale (Fig. 2.7), Nectaris, and even Schrödinger (Fig. 8.8b) are incised by these troughs. Originally thought to be tectonic features because of their often impressively straight walls, it is now believed that they are created by lines of coalescing secondary craters in a manner similar to that of "Imbrium Sculpture," although on a much larger

scale. The reason for the precise alignment of these ejecta clusters is unknown, although they are clearly common.

Concentric ridges and troughs occur near the inner rim of both large craters and basins, although they are seldom prominent on the moon and are frequently obliterated there by later mare infilling. The inner regions of basin ejecta blankets are also characterized by knobby or fractured and crenulated facies that probably reflect a high proportion of impact melt. Cooling and contraction of thick melt sheets is thought to produce the randomly oriented networks of fractures observed on the surfaces of these deposits.

Fragmental flow fronts and debris streams occur in the distal parts of basin ejecta blankets. These debris streams are usually longitudinally (radially) striated, but they are frequently deflected by topographic gradients (Fig. 6.6). When the debris streams run into topographic obstacles, transverse ridges called *deceleration dunes* develop. These features are evidently created by a ground-hugging flow of ejected debris, consistent with the concept of ballistic sedimentation.

Smooth plains fill topographic lows in and beyond the zone of debris streams. The origin of these plains is controversial. Such units are globally distributed around the moon. Originally called the "Cayley Formation" and thought to be volcanic, when sampled by Apollo 16 the smooth plains were shown to be underlain by fragmental material that is somehow related to basin formation. Orbital geochemical investigations show that much of this material is locally derived. Perhaps the smooth plains develop when high-velocity secondaries strike adjacent slopes, mobilizing the material on them, or perhaps seismic shaking from the impact causes large avalanches. Recent reinterpretations suggest that the "Cayley Formation" deposits are lava flows that are only thinly mantled by local ejecta. More sample returns and in-situ studies are probably required before the origin of these enigmatic smooth plains is understood.

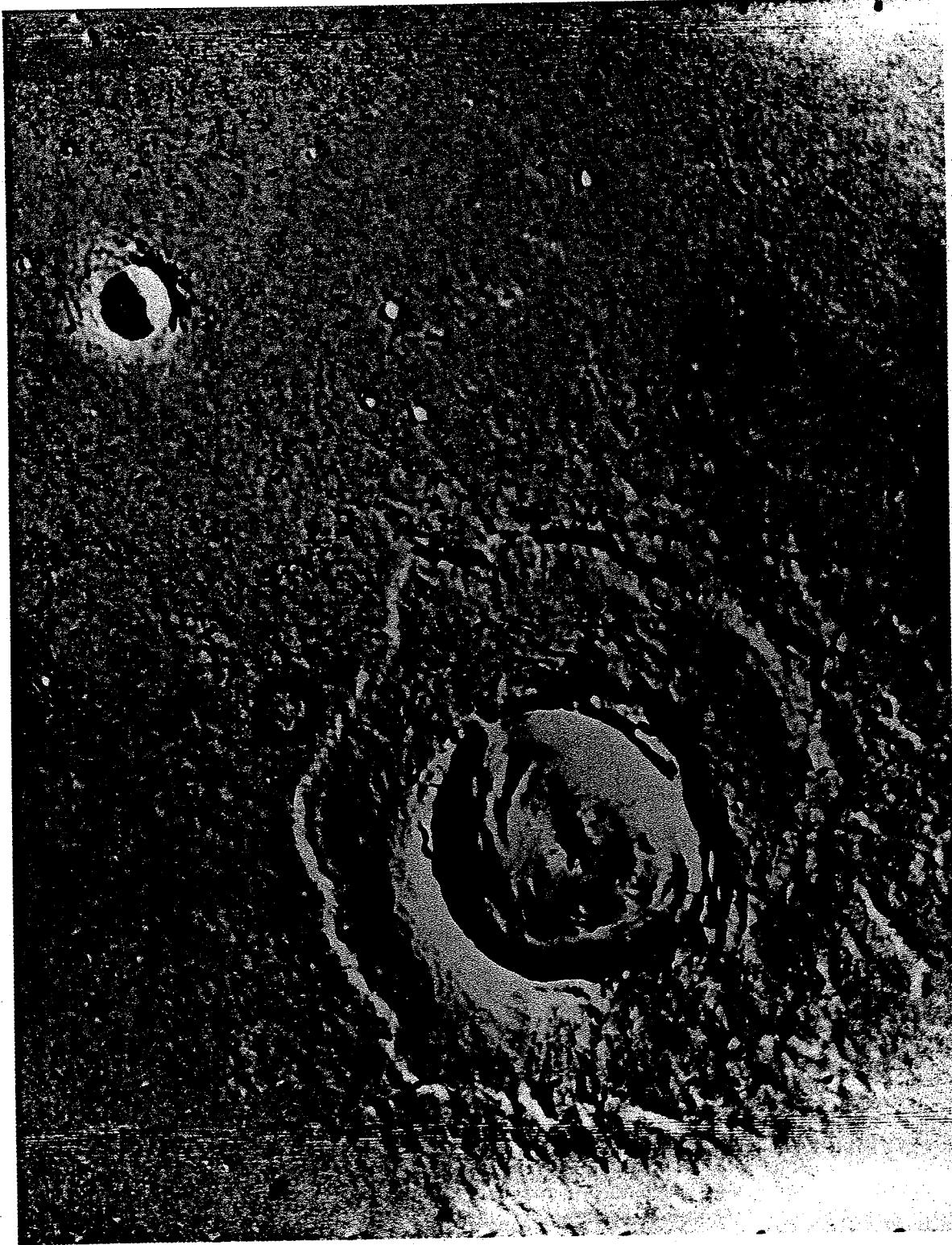
### 6.3.5 Fluidized ejecta

The ejecta blankets of impact craters on Mars are distinctly different from those on the moon or Mercury. Martian craters smaller than about 5 km in diameter closely resemble their counterparts on the moon. However, craters between 5 and 15 km in diameter have a single ejecta sheet that extends about one crater radius from the rim and ends in a low concentric ridge or outward-facing escarpment (Fig. 6.7a). These are called *rampart craters* from the continuous ridge surrounding the ejecta deposit. The ejecta of most craters larger than 15 km in diameter are divided into petal-like lobes that extend two or more radii from the rim (Figs. 6.7b and 6.7c), approximately



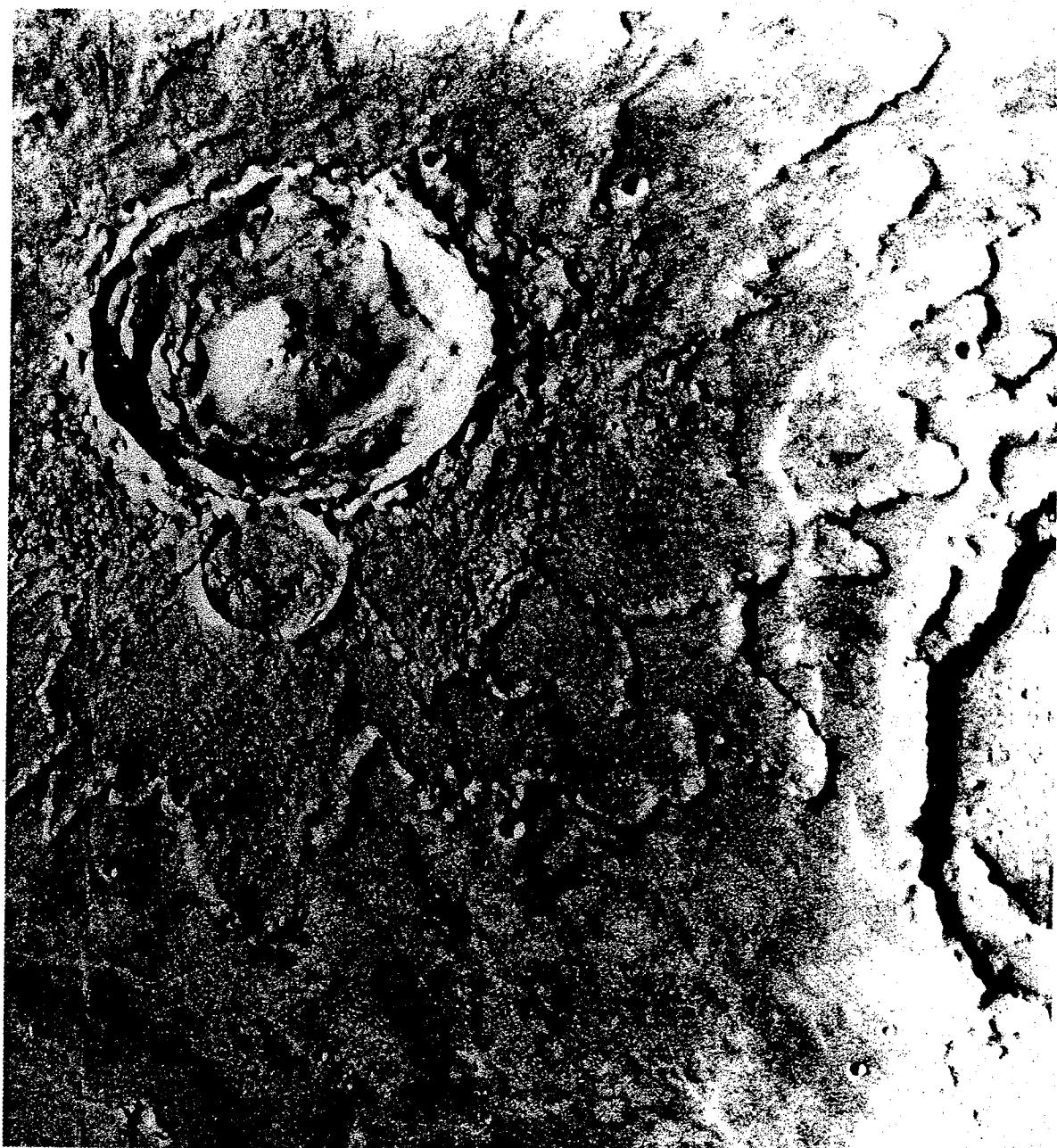
**Fig. 6.6** The distal portion of the Orientale basin ejecta blanket, about 1100 km southeast of the basin's center, shows evidence of a large-scale ground-hugging debris flow. Debris streams traverse diagonally across this photograph from upper left (toward the source) to lower right. Deceleration dunes are apparent near the left side of the photograph and debris has flowed into and ponded within a 90-km diameter crater (Inghirami) at the lower left. The horizontal lines across the photograph are 13 km apart. *Lunar Orbiter photograph LO IV 172 H2.*

(a)



**Fig. 6.7** Fluidized ejecta blankets on Mars. (a) A 9-km diameter rampart crater in Chryse Planitia. The ejecta blanket is clearly terminated by a low ridge or escarpment. Faint secondary craters and radial striae occur at larger distances. *Viking Orbiter frame 10A56.* (b) The 19-km diameter crater Yuty is surrounded by thin, petallike flow lobes that extend approximately twice as far from the crater as the continuous ejecta deposits of lunar or Mercurian craters. *Viking Orbiter frame 3A07.* (c) Martian crater Arandas is 28 km in diameter. Its ejecta blanket forms a circular plateau that ends in a convex-upward, outward-facing escarpment. Thin, radially lined flow lobes extend well beyond this escarpment. *Viking Orbiter frame 32A20.*

(b)



twice as far as the continuous ejecta deposits of lunar or Mercurian craters. A few large craters with lunar-type ejecta blankets are known, but they are rare. Many Martian craters have abnormally large central peaks and other internal collapse structures compared with lunar or Mercurian craters, also suggesting the presence of some fluidizing agent peculiar to Mars.

The lobate ejecta patterns of Martian craters larger than 15 km in diameter have been divided into two main types. The first is illustrated by the 19-km diameter crater Yuty (Fig. 6.7b). Many thin, radially lineated flow lobes extend away

from a narrow hummocky zone near the rim. Each flow lobe ends in a rampart. The second type, illustrated by the 28-km diameter crater Arandas (Fig. 6.7c), has a thicker, more circular ejecta deposit that ends in a convex outward-facing escarpment. Thin radially lineated flow lobes can sometimes be seen beyond the edge of the thick ejecta deposit. Many other Martian crater forms, such as pedestal craters, are probably the result of modification by erosional processes and will therefore not be treated here.

The fluidized ejecta blankets of Martian craters appear to have been emplaced as thin ground-

(c)

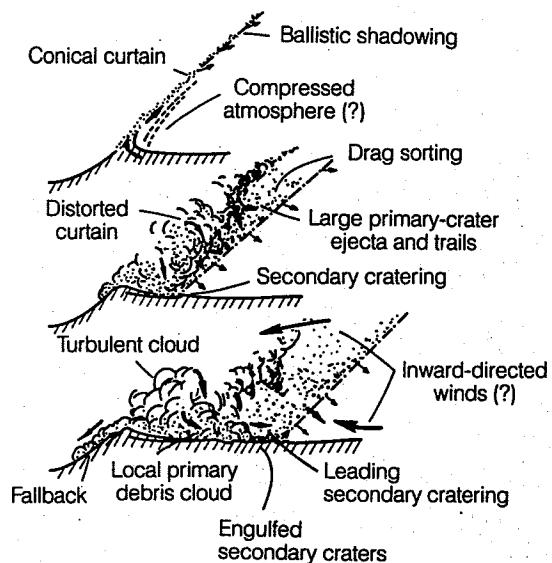


hugging flows. When impeded by topographic obstacles that could not be overridden, the flows are deflected and either spread out elsewhere or pond against the obstacle. Ejecta lobes fail to overtop low hills and mesas that are only a few times higher than the flow thickness itself, suggesting that the lobes could not have traveled as dispersed clouds of the base-surge type, nor were they emplaced by ballistic sedimentation, because the ejecta curtain should have fallen on the topographic obstacle from above.

The peculiar form of Martian ejecta blankets is generally attributed to the presence of liquid water in the substrate. Craters smaller than 5 km in diameter are supposed not to intersect the water table and so do not have fluidized ejecta blankets. Since 5-km diameter craters eject material from about one-third of their 2-km transient depth, this implies the existence of liquid water at depths of 600 to 1000 m over large portions of Mars. Frozen water near the surface cannot account for ejecta fluidization because most of the ejected material is not shocked to high enough pressures to melt ice (this requires about 10 GPa, which is reached only locally near the impact site).

Liquid water in the ejecta would greatly enhance the mobility of the ejected debris, converting the dry fragmental ejecta flows characteristic of lunar craters to fluid debris flows similar to terrestrial mudflows. In analogy with mudflows, the fluidity of the debris mass should depend upon both the content of water and of clay-size fines. Water-fluidized ejecta flows would naturally extend to a greater range from the crater rim, as observed, and in the process become much thinner than their lunar counterparts. The terminal ramparts of Martian crater ejecta appear similar to the terminal ridges of terrestrial mudflows, which are often armored by large boulders. Laboratory experiments by D. E. Gault and R. Greeley (1978) on impacts into mud targets have also produced ejecta forms similar to those of Martian craters.

An alternative suggestion was made by P. H. Schultz and D. E. Gault in 1979, who performed theoretical and experimental studies of the effects of an atmosphere on the expanding ejecta curtain. Their results, summarized in Figure 6.8, show that the largest ejecta are not significantly affected by atmospheric drag and continue to travel along the trace of the classic ejecta curtain, driving a high-speed outward-directed wind of displaced atmosphere ahead of them. Smaller ejecta, however, are winnowed from the curtain and either fall short of their ballistic range or, for the finest ejecta, form a turbulent cloud that deposits fragmental debris by a base-surge-type density current. Inward-directed winds may occur in the later stages as atmospheric gases rush in to fill the partial vacuum created behind the expanding ejecta curtain.



**Fig. 6.8** The effect of Mars' atmosphere on impact crater ejecta. Atmospheric drag winnows fine ejecta from the expanding curtain. The resulting turbulent dust cloud over the crater falls back producing density currents (base surges) whose deposits might explain some features of Martian fluidized ejecta blankets, according to the hypotheses of P. H. Schultz and D. E. Gault. Larger ejecta fragments are not deflected and produce secondary craters. Winds driven by the expanding ejecta curtain may modify the deposition of ejecta. *After Schultz and Gault (1979).*

Atmospheric modification of the ejecta curtain may explain the ejecta deposits of rampart craters, but even Schultz and Gault recognize that the multiple flow lobes of craters in the 15-km-and-up diameter range must be due to the incorporation of liquid water in the ejecta. If rampart craters are created by atmospheric modification, however, the inferred depth to liquid water is deeper than the 600 to 1000 m estimated previously. Supposing that abundant liquid water is required only for multilobed craters, liquid water occurs deeper than 2 to 4 km.

The atmosphere may also play a role in suppressing secondary craters around Martian primary impacts. Although secondary crater fields are known, they are seldom observed around craters with fluidized ejecta. It appears that whatever processes lead to fluidized ejecta patterns also suppress the ejection of large secondary crater-forming fragments.

In summary, the most straightforward explanation of the Martian multilobed ejecta patterns is that liquid water is incorporated into the ejected debris. Much research must be done, however, before this conclusion is secure and many details of the cratering process in water-saturated targets await clarification.

### 6.3.6 Oblique impact

Impact angle has a far more profound effect upon the ejecta blanket than it does upon the crater itself. The crater rim remains circular down to an impact angle of 10 degrees or less, whereas the ejecta blanket is noticeably modified at angles as high as 60 degrees. As the impact angle decreases from 90 degrees, the ejecta first show a preferential concentration on the downrange side of the crater. At incidence angles less than 45 degrees, a wedge-shaped "forbidden zone" (a term introduced by D. E. Gault and J. A. Wedekind, 1978) develops uprange of the crater (see Figure 6.9a). This wedge expands as the angle decreases until at angles less than about 20 degrees a second forbidden zone develops downrange of the crater (see Figure 6.9b). This second zone may be crossed by bright rays or streaks emanating from the elongated crater.

The ejecta deposits of oblique impacts are always bilaterally symmetric and frequently permit determination of the direction of impact. The concentration of ejecta in a sidewise direction was first recognized and described as a "butterfly" pattern by D. E. Gault and J. A. Wedekind (1978). It is evidently created by the elongated canoe-shaped shock wave produced by low-angle impactors. This shock wave mainly expands sidewise and throws the ejecta out away from its line of advance. The effect is rather similar to that of a plow in concentrating ejecta perpendicular to the line of travel of the projectile.

The precise angles at which the forbidden zones develop in the ejecta patterns are not well known. Most of our present knowledge rests on a few experiments performed at less than 6 km/second impact velocity at the NASA Ames Vertical Gun Facility into sand or pumice targets. Little theoretical understanding of these ejecta patterns yet exists. More work is needed, especially since ejecta patterns exhibiting the effects of obliquity are common, even for basin-sized impacts. A well-known example is the "butterfly" ejecta pattern of Orientale Basin on the moon, but many others may be seen around large craters on Mars.

### 6.3.7 Halo craters

Halo craters are not strictly ejecta facies, but they are so widely cited in the planetary geology literature that they deserve a brief mention here. "Dark halo craters" are simply craters whose ejecta deposit is darker than the surrounding terrain. Conversely, "light halo craters" have brighter ejecta than their surroundings. In both cases it is presumed that the impact excavates material of a different character than the surrounding surface, and so a halo is evidence for a buried geologic unit. Dark halo craters on the young bright terrain of Ganymede, for example, are cited as ev-

idence for the presence of old dark terrain some distance below the surface. Since impact craters excavate material from no greater depth than about one-tenth of their diameter, the thickness of geologic layers may be estimated from the minimum diameter at which halos form.

## 6.4 SECONDARY CRATER FIELDS

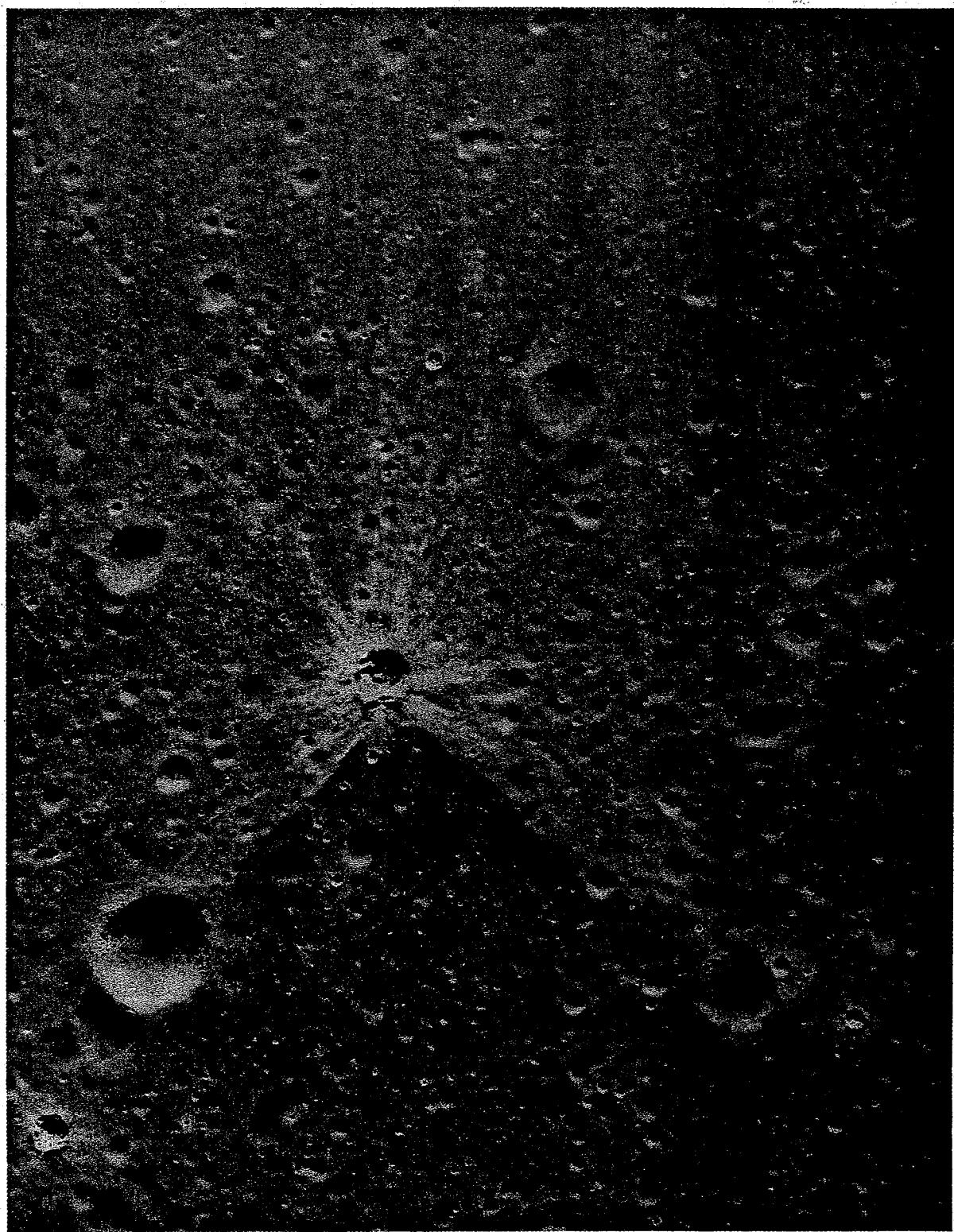
Large impact craters are surrounded by numerous secondary impact craters, variously occurring either singly or in loops, clusters, and lines. Figure 6.10 shows the famous secondary ejecta field surrounding Copernicus Crater on the moon. The secondary craters close to the primary one are irregular in planform and in depth, partly due to interference from other adjacent secondaries and partly because of the relatively low impact velocity of the fragments that produce the craters. Secondary craters are typically asymmetric, having steeper slopes on the side closest to the primary crater. As the range from the primary crater increases, the secondaries become more circular and clusters become more widely dispersed.

Secondary craters are first recognizable beyond the edge of the continuous ejecta blanket. From there, they extend many crater radii from their primaries. Indeed, the outer "edge" of a secondary crater field is undefinable because as the density of secondaries declines they become progressively more difficult to distinguish from the background primary crater population. The size of the largest secondary crater is roughly proportional to the size of the primary crater (the largest secondary crater diameter is about 4 percent of the primary diameter on the moon), so that the secondaries of basin-forming impacts may approach several tens of kilometers in diameter. The range of this largest secondary is also roughly proportional to the primary size. As a result, it is sometimes difficult to tell whether a given crater is primary or whether it is a distant secondary of a much larger event. This phenomenon causes some difficulty in inferring the primary impactor flux from the crater population on a given planetary surface.

One feature that often aids in distinguishing secondary from primary craters is the presence of a low chevron-shaped dune lying between two adjacent secondaries. Such dunes are apparent in Figures 6.2 and 6.10. When they form between craters in a chain the outward-pointing legs of the chevrons give the appearance of a fish skeleton, hence the common name "herringbone pattern."

The cause of the herringbone pattern was investigated experimentally by V. R. Oberbeck and R. H. Morrison in 1974. They found that the dunes develop when the expanding ejecta curtains of the adjacent secondary craters interfere with one another. Since secondaries strike at nearly the same time, the ejecta curtain of the first crater to

(a)



(b)

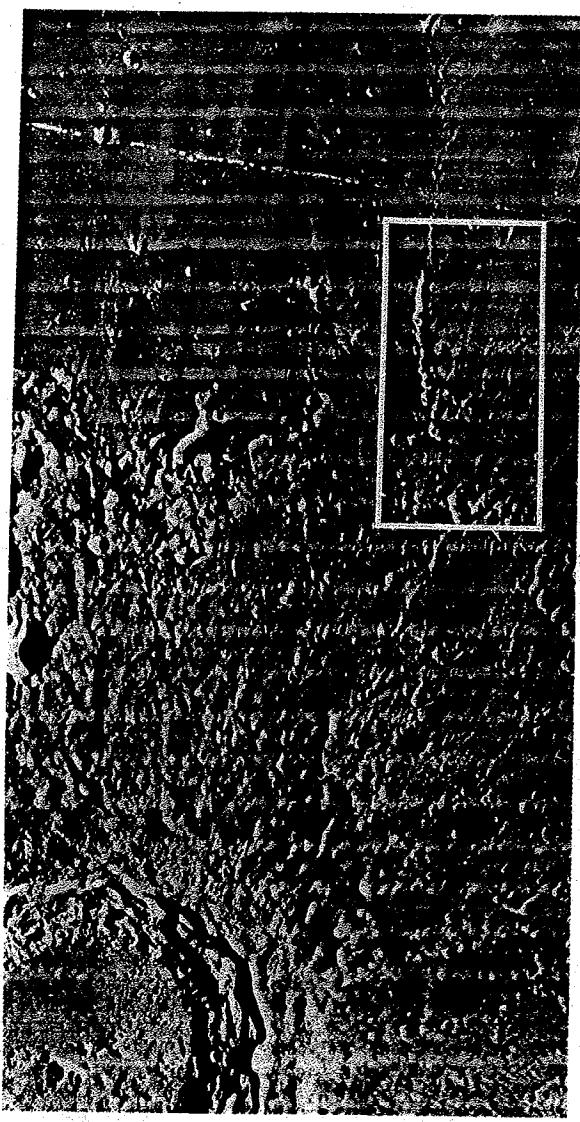


**Fig. 6.9** Asymmetric ejecta patterns produced by oblique impacts. (a) Typical pattern for impact angles between 20 and 45 degrees. The crater itself is still circular at this angle, but a triangular zone free of ejecta forms uprange of the projectile, which thus traveled from left to right in this photograph. *Apollo photograph AS15-9337.* (b) Very low impact angles (probably less than 6° for the crater in this photograph) create elliptical craters as well as very asymmetric “butterfly wing” ejecta patterns that often show downrange streamers (not present in this example). *Apollo photograph AS15-9254.*

(a)



(b)



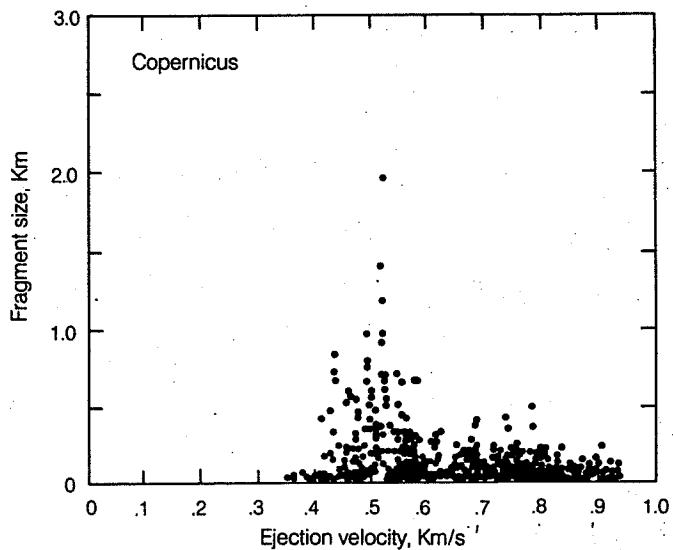
**Fig. 6.10** The secondary crater field surrounding the 93-km diameter lunar crater Copernicus. Secondary craters occur in loops, lines, and clusters near their primary, but become more separated at greater ranges. The inset shows a detail of a line of secondary craters on which prominent "herringbone" structures have developed. These structures are believed to originate from the interference of the secondaries' ejecta curtains. The inset also shows the typically irregular form of the secondary craters. Higher velocity secondaries at greater ranges are more bowl shaped and may be difficult to distinguish from small primaries. *Lunar Orbiter photograph LO IV H121. Inset LO V M-144.*

form is still expanding ballistically when the next crater is excavated. The ejecta curtain of this crater runs into that of the first. Many ejecta fragments collide and fall nearly vertically to create a dune between the craters. Because the first crater's ejecta curtain has more time to expand than that of the second, the dunes are deflected into a "V" enclosing the second crater to form, resulting in the observed chevron pattern. The orientation of the chevron thus indicates the order of formation of the adjacent secondary craters. If two equal-size craters form at precisely the same time, the

"chevron" dune is actually straight, bisecting the line between the centers of the secondaries.

Since it is highly unlikely that two independent primaries would strike so close together in space and time as to interfere with each other's ejecta curtains, the presence of chevron or herringbone dunes between craters is diagnostic of their secondary origin. This is sometimes useful for distant secondaries, which may not betray a secondary origin from their shapes but between which chevron dunes occur.

The block sizes inferred from secondary crater



**Fig. 6.11** Fragment sizes as a function of ejection velocity inferred from secondary crater diameters and ranges from lunar crater Copernicus. Fragments as large as 1 km in diameter were ejected at velocities near 0.5 km/second. The fragment size decreases with increasing velocity out to almost 1 km/second, where the fragment diameter abruptly decreases to below the limit of resolution. (*Courtesy A. M. Vickery.*)

diameters are often prodigious. Figure 6.11 shows that several blocks greater than 1 km in diameter were ejected from Copernicus at speeds exceeding 0.5 km/second. Studies of secondary crater fields on the moon, Mercury, and Mars show that the maximum fragment size is inversely proportional to the ejection velocity. The 20- to 30-km-diameter secondary craters associated with basin-forming impacts must have been produced by blocks (or, perhaps, by tight clusters of fragments) ranging from 3 to 4 km in diameter that were ejected at speeds of several kilometers per second. These enormous block sizes are a surprise because joints, layers, and other geologic discontinuities occur on a much smaller distance scale. Perhaps the inferred "blocks" are not coherent blocks at all, but are merely loose clusters of smaller fragments that are ejected together and follow nearly the same trajectory. The shallow, irregular morphology of secondary craters within a few primary crater radii of the rim closely resembles laboratory experiments performed on clustered impactors in which the individual fragments are too close together to produce separate craters. In my opinion, however, this explanation is more difficult to accept for the more distant secondaries that often show a normal bowl-shaped morphology.

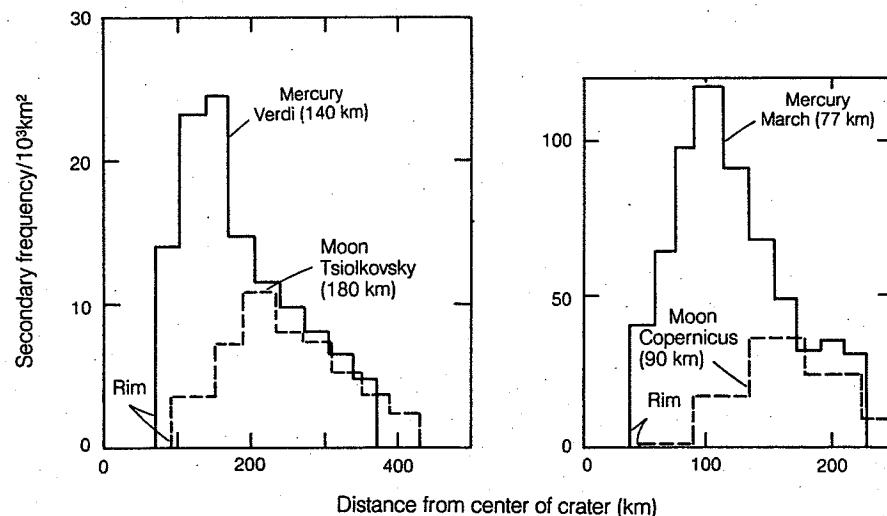
The inferred secondary block sizes are often much larger than the maximum size of blocks observed on the crater rim. Although this observation has been used as an argument against the reality of the inferred block size, the mechanics of emplacement of the rim blocks is much different from the mechanics of secondary fragment ejec-

tion and no valid comparison is possible. Rim blocks are ejected from some depth below the preimpact surface and thrown onto the rim as part of the highly sheared excavation flow. Secondary fragments are probably ejected from the near-surface spall zone by stress-wave interference and suffer little interaction after their ejection. They may thus attain larger sizes than the rim blocks. See the discussion of fragment sizes below.

Intact "megablocks" of Malm limestone have been found great distances from the 22-km diameter Ries Crater in Germany. These blocks, all from the uppermost geologic unit in the target area, range from one block 1000 m in diameter that was thrown at least 7 km from the transient rim down to blocks 25 m in diameter at ranges of 35 km, corresponding to an ejection velocity of 600 m/second. The existence of such large ejected blocks on earth suggests that the block sizes inferred from extraterrestrial secondary craters are real.

The extent of the secondary crater field is evidently a strong function of gravity. Even casual inspection of Mariner 10 images shows that Mercurian secondary crater fields lie proportionately closer to their primaries than their lunar counterparts (for example, compare Figures 6.10 and 8.8d). The secondary crater density within the field is also higher on Mercury than on the moon. This relation is quantified in Figure 6.12, where secondary crater density (for secondaries larger than an arbitrary lower limit of 1.4 percent of the primary diameter) is plotted against distance from the crater center for craters of comparable

## IMPACT CRATERING



**Fig. 6.12** The secondary crater density is compared around primary craters on the Moon and Mercury. Secondary craters on Mercury, which has approximately twice the moon's surface gravity, lie closer to their primary and are much more densely packed than on the moon. The craters chosen for comparison in each frame are related, as closely as possible, by the  $g^{1/6}$  scaling relation expected for large craters (Sec. 7.8). After Gault *et al.* (1975).

size on the moon and Mercury. Evidently, Mercury's two times greater gravity restricts the secondary crater range to roughly half the lunar value and raises the crater density to nearly four times that of the moon. A similar restriction of the secondary crater field has also been observed on Mars. On earth, secondary crater fields should

be even more closely confined to the vicinity of the primary crater and the secondary crater density should be correspondingly greater.

The quantity of material that produces the secondary crater field is a very small fraction of the total ejected. Table 6.1 lists the ratio between the total volume of the blocks that produced second-

TABLE 6.1 Volume of secondary crater-producing fragments

Primary Crater Name	Diameter (km)	Estimated* Crater Volume ( $\text{km}^3$ )	Volume of Fragments That Form Secondaries† ( $\text{km}^3$ )	Secondary Fragment Volume	
				Crater Volume	
<i>Moon</i>					
Theophilus	100	$2.3 \times 10^4$	$1.4 \times 10^2$	.006	
Copernicus	93	$1.9 \times 10^4$	$4.4 \times 10^2$	.023	
Bullialdus	61	$5.2 \times 10^3$	$1.49 \times 10^1$	.003	
Harpalus	45	$2.1 \times 10^3$	$1.91 \times 10^1$	.009	
<i>Mercury</i>					
Mozart	225	$2.7 \times 10^5$	$3.6 \times 10^3$	.013	
Verdi	150	$7.9 \times 10^4$	$2.2 \times 10^3$	.028	
Zola	60	$5.0 \times 10^3$	$1.31 \times 10^2$	.026	
<i>Mars</i>					
Lyot‡	220	$2.5 \times 10^5$	$5.2 \times 10^2$	.002	
Holden‡	175	$1.3 \times 10^5$	$5.9 \times 10^2$	.005	
Xc (MC15SW)	29	$5.7 \times 10^2$	$4.28 \times 10^1$	.075	

\*Crater volume  $V$  is estimated by assuming an initial transient crater two-thirds the diameter  $D$  of the final crater with a depth/diameter ratio of 1/5. This yields a volume  $V = 2.33 \times 10^{-2} D^3 (\text{km}) \text{ km}^3$ .

†The diameter  $d$  of the fragment that produces a secondary crater of diameter  $D_c$  is computed from the scaling equation (7.7.8) for competent or saturated rock,  $d = 0.671 D_c^{1.28} (g/v_c)^{0.277}$  (SI units), where  $g$  is surface gravity and the secondary impact velocity  $v_c$  is computed from the measured distance of the secondary from the primary's rim using Equation 6.1.3. The sample for most craters listed contains several thousand secondaries. Data from Vickery (1987) and Vickery, 1987 personal communication.

‡The secondary crater fields of these primaries were probably exhumed by erosion, so the total volume of secondary fragments inferred from the presently exposed secondary craters may be much smaller than the actual volume.

ary craters and the crater volume for a number of different-diameter craters on the moon, Mercury, and Mars. Secondary craters typically account for between 1 percent and 3 percent of the material ejected from the crater. This is a great deal less than the approximately 30 percent that lies beyond the range of the continuous ejecta, according to Equation 6.3.3. Evidently, most of the ejecta in the "discontinuous" ejecta zone fails to produce secondary craters. It is perhaps not surprising that large masses of rock material ejected at speeds of kilometers per second should be reduced to fragments too small to produce observable secondary craters. This leaves unsolved, however, the problem of how some large blocks survive to produce the observed secondaries.

One explanation of this problem is that the secondary-producing fragments are ejected by a different process than the one that throws out the bulk of the ejecta blanket. Most of the ejecta is expelled by the subsonic cratering flow that develops out of the residual particle velocity left behind after the shock wave passes (see Chap. 5). A small quantity of material, however, is ejected directly by stress-wave interference in the near-surface zone (see Sec. 5.4). Although it is ejected at high speed, this material is only lightly shocked and is thus a possible source of any unusually large blocks. The quantity of spalled material is roughly a few projectile masses and is thus in fair agreement with the quantities inferred from the number, ranges, and sizes of secondary craters.

At low speeds (up to a few hundred meters per second) the near surface ejecta consists of spall plates. These plates are typically several to ten times broader than they are thick. Their thickness  $\ell_s$  is approximated by the equation

$$\ell_s = \frac{T}{\rho c_L v_e} L \quad (6.4.1)$$

where  $L$  is the projectile diameter,  $T$  is the dynamic tensile strength of the target rock,  $\rho$  is its density,  $c_L$  is its longitudinal sound speed, and  $v_e$  is the spall ejection velocity. Rock beneath the spall plate is broken into much smaller fragments whose mean size  $\ell_{GK}$  can be estimated from a fragmentation theory developed by D. E. Grady and M. E. Kipp in 1980:

$$\ell_{GK} = \frac{T}{\rho v_e^{2/3} v_i^{4/3}} L \quad (6.4.2)$$

where  $v_i$  is the impact velocity. At ejection velocities greater than about 500 m/second, the spall plates contain so much elastic energy from the interfering stress waves that they themselves break up into the smaller, Grady-Kipp fragments (Eq. 6.4.2) immediately after ejection.

The ejection of large, thin spall plates and their subsequent fragmentation may explain the observed tendency of secondary craters to form in

clusters, although the linear chains still require explanation. Plate fragmentation at velocities near 500 m/second also accounts for the rapid decline in fragment size apparent at higher velocities.

The spalled material is also generally the highest-velocity solid ejecta (only jetted material is faster, and that is usually liquid or vapor). Ejection velocities may approach the maximum particle velocity in the target, which, for projectile and target material of similar shock impedances, is close to half of the impact velocity itself. The mass  $m_{ej}$  of material ejected by the spall mechanism at velocities greater than  $v_e$  and which is shocked to pressures less than  $P_{max}$  is estimated in units of the projectile mass  $m$  by:

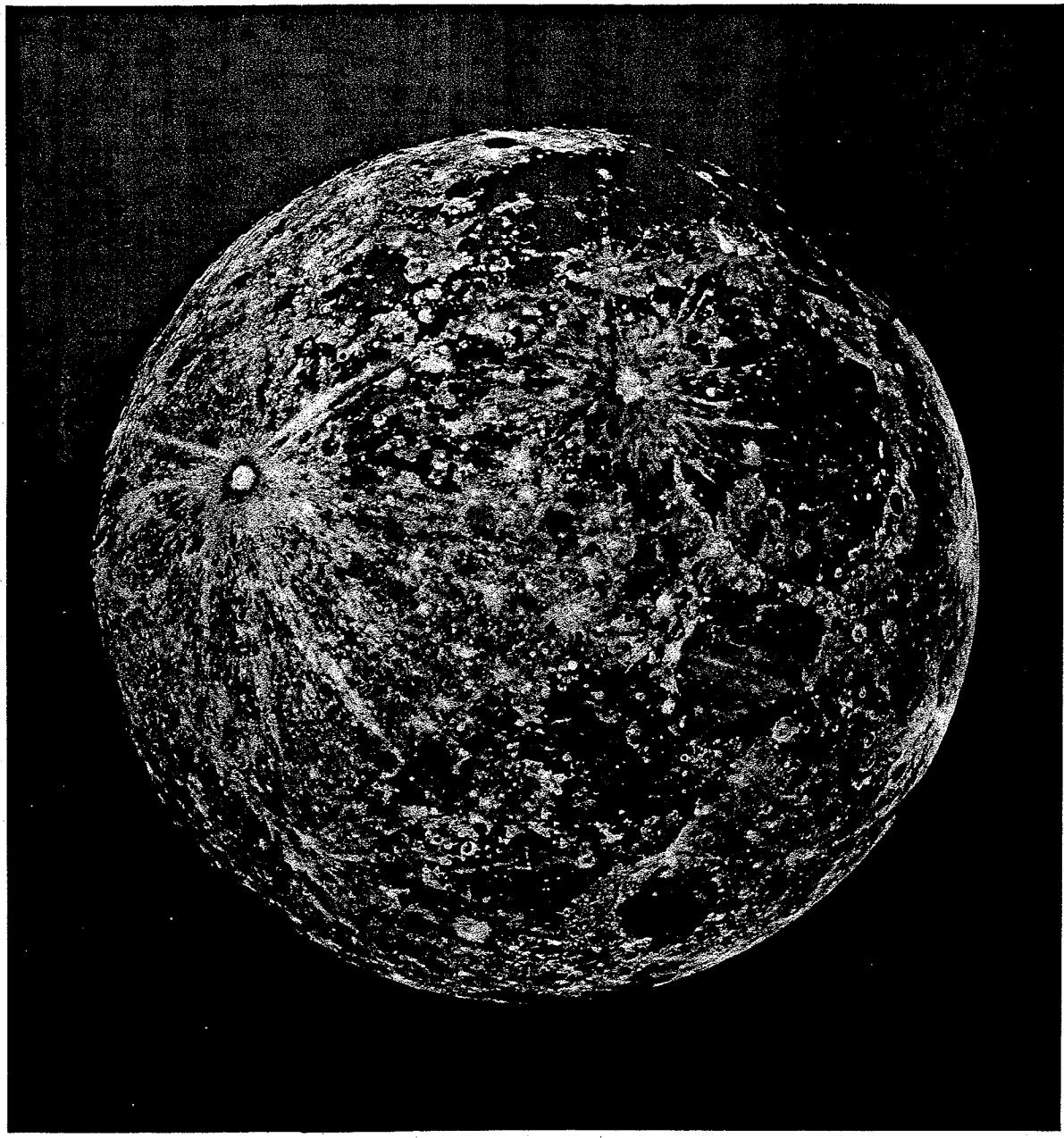
$$\frac{m_{ej}}{m} = 1.2 \frac{P_{max}}{\rho c_L v_i} [1 - (2v_e/v_i)^{1/3}] \quad (6.4.3)$$

Equation 6.4.3 indicates that a small quantity of solid material from the near-surface region may be ejected at speeds exceeding planetary escape velocity. This prediction was spectacularly confirmed in 1983 by the recognition of the first lunar meteorite ALHA 81005. This golfball-sized chunk of lunar breccia clearly lay on the lunar surface before it was ejected at more than 2.5 km/second, probably from a crater 3 km or more in diameter. In agreement with the spall model, the meteorite shows no petrologically detectable shock damage from the ejection event.

The growing conviction that the 8 SNC (Shergottite, Nakhelite, and Chassigny) meteorites originated on Mars also implies ejection at high speed and points to a spall-type mechanism, although in this case the relatively high shock levels recorded by the Shergottites and Chassigny indicate that they originated from greater depth in the spall zone than the lunar meteorites. Many other meteorites seem to have been ejected from the near-surface regions of their parent bodies without significant shock damage, making the spall mechanism more important scientifically than the relatively small amount of material ejected by this mechanism might suggest. The fact that the large, intact "megablocks" ejected from the Ries Crater originated from the uppermost geologic unit exposed at the impact site is also consistent with the supposition that most secondary craters are formed by fragments ejected by the spall process.

## 6.5 CRATER RAYS

Crater rays are the most distant and often the most conspicuous form of ejecta. Rays are bright albedo features that extend from fresh craters in subradially to radially oriented filaments and diffuse patches, often forming streaks that approximate great circles across the surface of the planet or satellite on which they form. The rays from Tycho (Fig. 6.13), for example, cross almost the entire visible hemisphere of the moon. Rays are

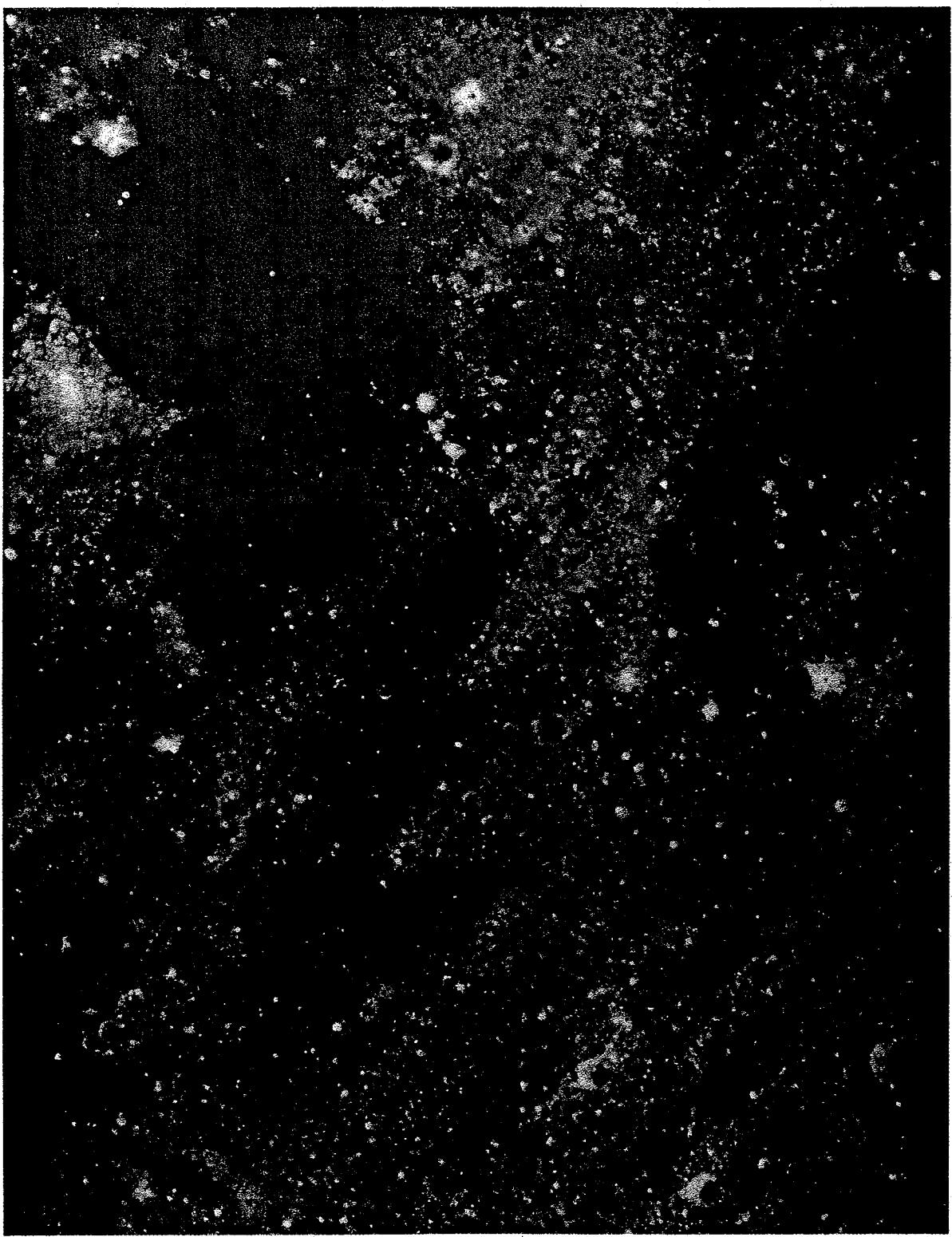


**Fig. 6.13** Rays from the young crater Tycho are visible over an entire lunar hemisphere. Rays are most prominent under full-moon (backscatter) illumination. The less extensive ray systems of Copernicus and Kepler appear at the upper right. Note the unconventional orientation: north is to the right. *Lick Observatory Photograph, courtesy E. Whittaker.*

brightest under high sun-angle illumination and may be readily seen at full moon when other topographic details are difficult to discern because of the lack of shadows. Rayed craters occur on all airless bodies that have been imaged to date. They have not been observed on Mars, either because the atmosphere interferes with their emplacement or because active aeolian processes erase them soon after formation.

Rays are ephemeral features. Of all fresh-appearing craters on the moon, for example, only

about 12 percent have recognizable rays, and a similar proportion holds for Mercury. It is believed that the micrometeoroid-induced "gardening" or overturn of the regolith eventually obliterates the rays, although the precise mechanism of obliteration is unknown. Rays on the lunar surface persist only 1 to  $2 \times 10^9$  years. Craters older than this lack rays, although they may still be classified as "fresh" on the basis of rim sharpness and general unmodified appearance. The rays of small craters probably fade faster than those of large



**Fig. 6.14** High resolution image of crater rays on the surface of western Mare Crisium. The bright rays are not obviously associated with small craters or any other surface structures. This frame is approximately 22 km across horizontally. *From the central portion of Apollo photograph AS15-9235.*

craters, but this cannot be well documented on the basis of available data. The presence of rays is commonly used in lunar stratigraphy to distinguish the Eratosthenian (rayless) system from the later Copernican (rayed) system, although this procedure may be invalid if the rate of ray disappearance depends upon crater size.

The maximum radius of the ray pattern is crudely proportional to crater size. R. B. Baldwin (1963) found, for example, that the radius of the ray pattern  $R_r$ , around lunar craters ranging between 3 and 200 km in diameter is a power function of crater radius  $R$ :

$$R_r = 10.5R^{1.25} \quad (6.5.1)$$

where all distances are measured in kilometers. Individual ray systems may, however, be either larger or smaller than this formula predicts by a factor of three.

The nature of rays is not well understood. At one time it was debated whether rays are mainly deposits of primary ejected material or whether the high albedos are related to secondary or even tertiary cratering processes. However, it now seems clear that the secondary impact velocity at ranges where rays form is so high that the primary ejecta must be mixed with a large fraction of substrate material. Although many rays are associated with resolvable secondary craters, others are not (Fig. 6.14), and it appears that secondary cratering on a very small scale can produce bright rays.

The reason that rays are bright is somewhat better understood. The low albedo typical of the lunar surface is now known to be due to the admixture of rock fragments and dark, impact-produced, glass-welded aggregates (agglutinates) in the regolith. Any increase in the abundance of crystalline silicates or fracturing of the glass-coated agglutinate particles leads to an increase in albedo. Ray-forming ejecta both add to the quantity of crystalline silicates and fracture the regolith particles that are already present, producing the observed bright streaks and patches. The high albedo rays fade as continued micrometeoroid impact on the regolith overturns the near-surface mixing layer (typically about 10 cm thick) and creates new agglutinates. Eventually, when the crystalline and glassy particle populations reach equilibrium, the ray becomes indistinguishable from the rest of the lunar surface.

The source of the ejecta that forms rays is poorly understood. It seems very unlikely that jetting is responsible for rays because shock levels in jetting are so high that most material is vaporized. Jetting also ejects material at speeds equal to or larger than the impact velocity so that it should all escape the planet or satellite. Rays may be created by high-speed debris ejected from the spall zone close to the projectile, which is crushed to

small fragments by large stress gradients. This could explain why secondary craters are not always associated with rays, since highly dispersed solid ejecta may still cause extensive fracturing in the near-surface regolith-mixing layer without producing recognizable pits. However, spectrophotometric evidence has been found for a small amount of lunar highlands material in a Copernicus ray that overlies mare basalt. The material must have come from beneath the Imbrium basalt penetrated by the projectile that formed Copernicus. This argues that ray-forming material does not originate near the surface but instead is thrown out as part of the deeper-seated excavation flow—an interpretation that is also consistent with the identification of dark, ropy glass in Apollo 12 soil collections as Copernicus ray ejecta. Although the glass in this case would not have caused an albedo increase, associated lightly shocked ejecta fragments could produce the observed ray's brightening. Such an association of lightly and highly shocked debris is characteristic of excavation flow processes (Sec. 5.5.2).

Rays also form in the ejecta of near-surface explosions. In a series of small-scale explosion cratering experiments, R. J. Andrews (1977) was able to correlate jets of high-velocity gas in the fireball with rays in the ejecta pattern. The formation of rays in these experiments is thus due to the expanding detonation products. Perhaps vapor expanding from the site of an impact plays an analogous role in the production of crater rays.

The number and length of the rays in Andrews' experiments are strong functions of the depth of burst. The number of rays increases rapidly as depth of burst decreases, while the length decreases. The reasons for these correlations are presently unclear.

Much remains to be learned about the nature and origin of crater rays. Since they seem to involve very little material, in situ investigation is probably necessary before much progress can be made.

## FURTHER READING

### *General References*

The literature on ejecta deposits is so scattered among different subfields that it is difficult to recommend any one general overview. The best general discussion is very out-of-date, but because of its historical importance I give the reference:

R. B. Baldwin, *The Measure of the Moon*, University of Chicago Press, Chicago, 488 pp. (1963).

Another important reference of similar vintage, also relying exclusively on earth-based data, ushers in the modern era of impact crater study:

E. M. Shoemaker, Interpretation of lunar craters, in Z. Kopal (Ed.), *Physics and Astron-*

*omy of the Moon*, Academic Press, New York and London, pp. 283-357 (1963).

Much of the more modern thinking on ejecta blankets is summarized in the context of a detailed study of the ejecta blanket of the terrestrial Ries Crater:

F. Hörz, R. Ostertag, and D. A. Rainey, Bunte breccia of the Ries: Continuous deposits of large impact craters, *Rev. Geophys. Space Phys.* 21: 1667-1725 (1983).

An important experimental investigation of impact ejecta, with emphasis on the provenance and shock level of material in the ejecta blanket is:

D. Stöffler, D. E. Gault, J. Wedekind, and G. Polkowski, Experimental hypervelocity impact into quartz sand: Distribution and shock metamorphism of ejecta, *J. Geophys. Res.* 80: 4062-4077 (1975).

### Special Topics

The classic paper on crater rims, including the important discovery of the overturned flap, is by E. M. Shoemaker (1963), cited at the end of Chapter 4. A more recent discussion of the topography of lunar crater rims is

M. Settle and J. W. Head, Radial variation of lunar crater rim topography, *Icarus* 31: 123-135 (1977).

The thickness of the ejecta deposits of lunar craters has been the subject of an extended debate. In addition to the above paper, the principal papers in this debate are

N. M. Short and M. L. Forman, Thickness of impact crater ejecta on the lunar surface, *Modern Geology* 3: 61-91 (1972).

T. R. McGetchin, M. Settle, and J. W. Head, Radial thickness variation in impact crater ejecta: implications for lunar basin deposits, *Earth Planet. Sci. Lett.* 20: 226-236 (1973).

M. Settle, J. W. Head, and T. R. McGetchin, Ejecta from large craters on the moon: discussion, *Earth Planet. Sci. Lett.* 23: 271-274 (1974).

The debate over ejecta deposit thickness and the understanding of flow features in these deposits was given a new direction by the following landmark paper:

V. R. Oberbeck, The role of ballistic erosion and sedimentation in lunar stratigraphy, *Rev. Geophys. and Space Phys.* 13: 337-362 (1975).

Facies in the ejecta blankets of lunar craters are clearly and briefly summarized by

K. A. Howard, Fresh lunar impact craters: Review of variations with size, in *Proc. Fifth Lunar Conf.*, vol. 1, pp. 61-69 (1974).

The best description of facies in impact basin deposits is

H. J. Moore, C. A. Hodges, and D. H. Scott, Multi-ringed basins—illustrated by Orientale and associated features, in *Proc. Fifth Lunar Conf.*, vol. 1, pp. 71-100 (1974).

There is much speculation about the origin of Martian fluidized ejecta blankets. Since this issue is far from settled, I merely cite three of the earliest and most important papers in the debate. This class of ejecta deposit was first reported by

M. H. Carr, L. S. Crumpler, J. A. Cutts, R. Greeley, J. E. Guest, and H. Masursky, Martian impact craters and emplacement by surface flow, *J. Geophys. Res.* 82: 4055-4065 (1977).

Rampart craters were produced in small-scale experiments by

D. E. Gault and R. Greeley, Exploratory experiments of impact craters formed in viscous-liquid targets: Analog for Martian Rampart Craters?, *Icarus* 34: 486-495.

An alternative to liquid water in the Martian subsurface was cogently argued by

P. H. Schultz and D. E. Gault, Atmospheric effects on Martian ejecta emplacement, *J. Geophys. Res.* 84: 7669-7687 (1979).

Most of the work on oblique impact was done by D. E. Gault and J. A. Wedekind, cited at the end of Chapter 7. Secondary crater fields were first studied in detail by E. M. Shoemaker, cited in the General References above. The "herringbone" pattern was extensively studied by

V. R. Oberbeck and R. H. Morrison, Laboratory simulation of the herringbone pattern associated with lunar secondary crater chains, in *The Moon* 9: 415-455 (1974).

A recent study of the fragment sizes that create the observed secondary craters is

A. M. Vickery, Size-velocity distribution of large ejecta fragments, *Icarus* 67: 224-236 (1986).

The process of spallation and the origin of the lunar meteorites are discussed by

H. J. Melosh, Ejection of rock fragments from planetary bodies, *Geology* 13: 144-148 (1985).

There are few comprehensive discussions of crater rays in the literature. One of the best is

C. C. Allen, Rayed craters on the moon and Mercury, *Phys. Earth Planet. Int.* 15: 179-188 (1977).

## VII. SCALING OF CRATER DIMENSIONS

### 7.1 INTRODUCTION

One of the most frequently asked questions about an impact crater is "how big was the meteorite that made the crater?". Like many simple questions this has no simple answer. It should be obvious that the crater size depends upon both the meteorite's size, speed, and the angle at which it struck. It also depends upon factors such as the meteorite's composition, the material and structure of the surface in which the crater forms, and the surface gravity of the target planet. The question about the original size of the meteorite is usually unanswerable because the speed and angle of impact are seldom known. However, many important scientific questions turn on the ability to estimate impactor size from the impact crater, given appropriate estimates of speed and other factors.

The inverse question, of how large a crater will be produced by a given size impactor with a known speed and incidence angle, is in principle much simpler to answer. However, even this prediction is fraught with uncertainty. Essentially, the problem is that there is no observational or experimental data on impact craters larger than a few tens of meters in diameter, while the impact structures of geologic interest range up to 1000 km in diameter. The traditional escape from this difficulty is to extrapolate beyond experimental knowledge by means of more-or-less well-founded *scaling laws*.

In this chapter "crater diameter" should be understood to mean the diameter of the "transient crater" that forms at the end of the excavation stage. Subsequent gravity-driven collapse, treated in Chapter 8, may enlarge this diameter by roughly 20 percent for small, bowl-shaped simple craters or by as much as 30 to 70 percent for the larger, more thoroughly collapsed complex craters. Nearly all experimental investigations report the transient crater's "apparent" diameter and depth: that is, the diameter at the level of the preexisting ground surface and the crater's maximum depth below this level. Apparent diameters are difficult to measure on extraterrestrial craters, so most photogeologic studies of craters report the rim-to-rim diameter. In this chapter the apparent diameter and depth will be used, indicated by the subscript "at" where a distinction is important. Small-scale laboratory craters (which are essentially *transient* craters because they almost uni-

versally lack the simple crater's breccia lens) have rim-to-rim diameters about 25 percent larger than their apparent diameters.

Figure 7.1 illustrates the quantities important in impact crater scaling. These quantities fall into three groups: projectile variables such as density  $\rho_p$ , mass  $m$ , velocity  $v_i$ , diameter  $L$ , kinetic energy  $W$ , and angle of impact  $\theta$ ; crater descriptors, such as diameter  $D_{at}$ , depth  $H_{at}$ , volume  $V$  and rim height  $\delta$ ; and target variables, density  $\rho_t$ , yield stress  $Y$ , gravity  $g$ , and porosity  $\phi$ . Not all of these variables are independent: for example, projectile mass can be expressed in terms of diameter and density.

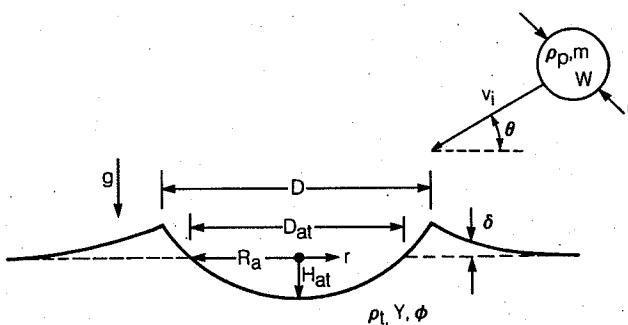
### 7.2 YIELD SCALING OF EXPLOSIONS

Scaling laws were first introduced in 1950 by C. W. Lampson, who studied the results of a series of TNT explosions using charges ranging from 4 to 1500 kg. Lampson's work was originally commissioned by the National Defense Research Committee at Princeton, N.J., in 1943-1945. Lampson noted that parameters describing different size explosions could be expressed by similar equations if all distances (e.g., depth, diameter) were divided by the cube root of the explosive energy (yield)  $W$ . The ratio between crater diameter and  $W^{1/3}$  was found to be nearly constant, independent of  $W$ . Crater diameter is thus proportional to  $W^{1/3}$ , so that, not surprisingly, larger quantities of explosive result in larger craters. This relation permits extrapolation of experimental results on small craters to predict the size of craters produced by larger explosions. Thus, if the diameter  $D_0$  of a crater produced by an explosive of energy  $W_0$  is known, and the diameter  $D$  of a crater produced by an explosive energy  $W$  is desired, the diameter  $D$  can be found from the following proportion:

$$\frac{D}{D_0} = \left( \frac{W}{W_0} \right)^{1/3} \quad (7.2.1)$$

An exactly similar proportion, or scaling law, may be written to relate the predicted depth  $H$  of a crater to the depth  $H_0$  of a measured crater.

Lampson's scaling law applies only to explosive charges placed at *similar* depths, that is, at depths that are related by the  $W^{1/3}$  scaling rule. Explosive charges with the same energy but placed at different depths produce different size craters, as shown



**Fig. 7.1** Variables and parameters important in impact crater scaling relations. These quantities are divided into three groups: Projectile variables are density  $\rho_p$ , diameter  $L$ , mass  $m$ , impact velocity  $v_i$ , kinetic energy  $W$ , and angle of impact  $\theta$ . Transient crater descriptors are apparent diameter  $D_{at}$ , radius  $R_a = D_{at}/2$ , distance from the crater's center  $r$ , crater depth  $H_{at}$ , crater volume  $V$  and ejecta blanket thickness  $\delta$ . Target variables are density  $\rho_t$ , yield strength  $Y$ , porosity  $\phi$ , and the acceleration of gravity  $g$ . Not all of these quantities are independent, but the different forms are used where convenient in this chapter.

in Figure 7.2. In this figure, *scaled* crater diameter is plotted against *scaled* depth of burial. Very shallow depths of burial result in small craters: crater size increases with the depth of burial until it reaches a maximum at the "optimum depth of burial," after which crater diameter decreases to zero for sufficiently deeply buried explosives.

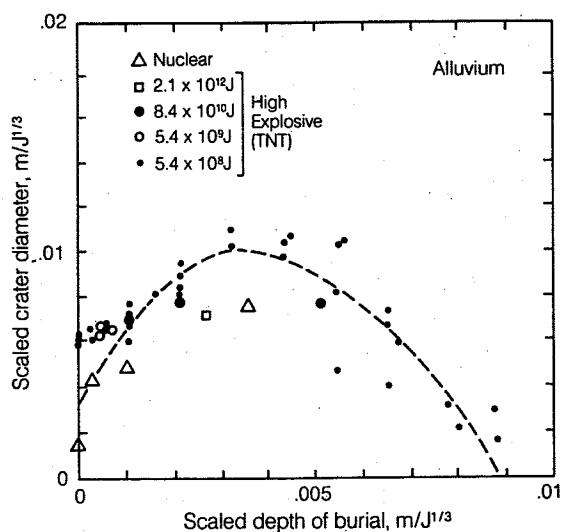
More recent work that incorporates data from large nuclear craters indicates that Lampson's 1/3 power law is not accurate over the entire range of crater sizes. A 1/3.4 law, also established empirically, works much better. Very large explosions, in which gravity plays a major role in limiting crater size, are expected to scale roughly as the power

1/4 if energy is the only important scaling variable (see Section 7.4). Nevertheless, the philosophy of scaling by a power of the explosive yield is the same, and reasonably accurate predictions for buried explosions are possible.

### 7.3 THE IMPACT-EXPLOSION ANALOGY

The development of scaling relations for explosion cratering was more than just an inspiration for impact cratering. Gifford, in 1924 and 1930, was one of the first to realize that the kinetic energy per unit mass in a meteoritic projectile far exceeds the energy density available in even the most powerful chemical explosives. Gifford correctly anticipated that a meteoritic impact is similar to an explosion. The modern view is expressed as the *impact-explosion analogy*, in which the impact of a body moving more than a few kilometers per second is likened to an explosion. Both impacts and explosions rapidly create a region with a large energy density near the surface of a planet. Once this energy is deposited, the subsequent expansion of the crater is similar in both cases. Equation 7.2.1 may thus be used for impacts, by substituting the kinetic energy of the impactor,  $\frac{1}{2}m v_i^2$ , for the explosive energy  $W$ . Impacts tend to deposit more momentum in the target than explosions, so the impact-explosion analogy should be regarded with some suspicion. Nevertheless, it provides some useful insights into the gross nature of the impact cratering process.

The major uncertainty in applying scaling laws developed for explosions to impacts is choosing a "depth of burial" for an impact. Although this quantity is well defined and easily measured for an explosion, the same is not true for impacts. Experience has shown that impacts are "shallow buried"; that is, their effective depth of burial is much less than the optimum depth of burial,



**Fig. 7.2** Explosion crater diameter vs. depth of burial in alluvium at the Nevada Test Site. This figure exemplifies  $W^{1/3}$  scaling and indicates the importance of the depth of burial for explosions. The equivalent depth of burial for impacts is close to zero. *After Nordyke (1962)*.

shown in Figure 7.2. Crater diameter is a steep function of the burial depth for shallow explosions, so serious errors can result from an incorrect choice.

Meteor Crater, Arizona, provides an excellent example. In 1960, E. M. Shoemaker estimated that the 1-km-diameter crater was excavated by a meteorite that deposited roughly 1.7 megatons ( $7.1 \times 10^{15}$  J) in the rocks of the Colorado Plateau. He obtained this number by assuming that the iron meteorite penetrated four to five times its own diameter and could thus be scaled from the deeply buried Teapot Ess nuclear explosion crater. In 1963, R. B. Baldwin estimated a penetration depth of only 0.9 projectile diameters and an energy of 8.1 megatons ( $3.4 \times 10^{16}$  J). Numerical modeling by J. B. Bryan and others (1978) suggested a penetration depth of two to three projectile diameters and an energy of 4.5 megatons ( $1.9 \times 10^{16}$  J). Still more recent studies (Schmidt, 1980) based on centrifuge scaling indicate an even shallower equivalent depth of burial and a formation energy between 10 and 60 megatons (8 to  $25 \times 10^{16}$  J), depending upon impact velocity. Obviously, the situation is still confused even for such a well-studied impact crater as Meteor Crater.

Until further work clarifies this situation, the equivalent depth of burial  $d_b$  for an impact may be roughly estimated from the classic jet-penetration formula

$$d_b \approx L \sqrt{\rho_p / \rho_t} \quad (7.3.1)$$

where  $L$  is the projectile diameter,  $\rho_p$  is its density and  $\rho_t$  is the target density. A number of experiments qualitatively support this equation. In particular, the prediction that dense projectiles penetrate more deeply and thus create larger craters than less dense projectiles of the same energy seems to be well verified. If Equation 7.3.1 is correct, then the scaled depth of burial ( $d_b / W^{1/3}$ ) decreases as impact velocity increases, which implies that crater size depends upon both projectile energy (through Lampson's scaling law, Equation 7.2.1) and indirectly upon impact velocity, since the scaled depth of burial is a function of  $v_i^{-2/3}$ .

The impact-explosion analogy is intuitively appealing and captures much of the essential nature of a high-velocity impact. However, it ignores the momentum deposited by the projectile in favor of its energy, the precise definition of the equivalent depth of burial is unclear, and it provides no information about what happens in an oblique impact. Furthermore, buried explosives produce large quantities of confined gas that play a major role in the early phases of crater excavation. Impact craters, in contrast, are vented from their inception. Until or unless a more precise statement of the impact-explosion analogy is found that in-

cludes the effects of these factors, its usefulness for quantitative purposes is limited.

#### 7.4 HYDRODYNAMIC SIMILARITY AND GRAVITY

A more mathematically rigorous form of scaling is derived from the equations of hydrodynamics. These equations are assumed to govern all macroscopic cratering events. This assumption is implicit when numerical computer codes are used in an attempt to model cratering processes, since these codes are, at best, merely approximate means of solving the hydrodynamic equations. Any scaling law valid for these equations will likewise be valid for impact cratering processes.

The full hydrodynamic equations are given below for reference. It is not necessary for the reader to know how to solve these equations. In fact, one of the powerful features of scaling is that it is possible to demonstrate the scaling properties of these equations without having to solve them.

The three equations of compressible fluid flow in the presence of a gravity vector  $g_i$  (where the index  $i$  ranges from 1 to 3 and denotes the  $x$ ,  $y$ , or  $z$  Cartesian components of the gravity vector) and deviatoric stress tensor  $\Sigma_{ij}$  are given by the mass conservation equation:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_i}{\partial x_i} = 0 \quad (7.4.1a)$$

Where a repeated index  $i$  or  $j$  implies a summation over the three vector components: for example,  $u_i u_i = u_1 u_1 + u_2 u_2 + u_3 u_3$ . The momentum conservation equation is:

$$\frac{\partial \rho u_i}{\partial t} + \frac{\partial}{\partial x_j} (\rho u_i u_j + P - \Sigma_{ij}) = g_i \rho \quad (7.4.1b)$$

And the energy conservation equation is:

$$\frac{\partial \rho E_i}{\partial t} + \frac{\partial}{\partial x_j} (\rho u_i E_i + P u_j - \Sigma_{ij} u_i) = \rho u_i g_j \quad (7.4.1c)$$

In this set of equations  $\rho$  is the material density (which may be a function of position and time),  $u_i$  are velocity vector components,  $t$  is time, and  $x_i$  are position vector components.  $P$  is the pressure, and the total specific energy density (per unit mass)  $E_i$  is given by

$$E_i = \frac{1}{2} u_i u_i + E \quad (7.4.2)$$

In addition to the preceding equations, a thermodynamic equation of state is needed, one that relates the pressure  $P$  to material density and internal energy density  $E$ :

$$P = P(\rho, E) \quad (7.4.3)$$

It is easy to show that most terms of these equations are invariant under a scale transformation

that preserves density unchanged and multiplies all distances and times by a constant factor  $\alpha$ :

$$\begin{aligned} t' &= \alpha t \\ x' &= \alpha x \\ \rho' &= \rho \end{aligned} \quad (7.4.4)$$

When this substitution is made, the new equations in  $x'$  and  $t'$  are essentially the same as the old equations in  $x$  and  $t$ . Velocities, pressures, and energy densities are unchanged. Only the time and distance scales are altered. This invariance is violated only by the gravity terms and (possibly) by the deviatoric stress  $\Sigma_{ij}$ .

The importance of this invariance is illustrated by an example: suppose that a laboratory experiment has been conducted on the impact of a 1-cm diameter iron sphere into a basalt target at 10 km/second and all of the quantities of interest, such as shock pressure, melt and vapor production, cratering flow velocity, and crater dimensions, have been measured. The validity of the hydrodynamic equations assures that the results of this experiment are identical to those of an impact involving, say, a 50-km diameter iron projectile striking a basalt target at the same speed, provided only that all distances in the laboratory experiment are multiplied by  $\alpha = 5 \times 10^6$ , all times are multiplied by the same factor, and densities are unchanged. Thus, an event that occurs 1 ms after impact in the laboratory will occur 5000 seconds (1.4 hours) after impact for the 50-km diameter projectile. Since the impact velocities are identical in the two impacts, shock pressures and energy densities are also identical at positions and times related by the factor  $\alpha$ . This scaling relation even holds in the presence of material strength, as long as the strength is unchanged by  $\alpha$  (that is,  $\Sigma'_{ij} = \Sigma_{ij}$ ). Although this is not true for rate-dependent processes, such as viscosity, it holds reasonably well for elasticity and strain-dependent plastic yield.

This result sounds almost too good to be true. Why, then, is there any difficulty predicting the results of planetary-scale impacts if valid scale-model experiments can be performed in the laboratory? The problem is the one term on the right-hand side of Equations 7.4.1b and 7.4.1c that does not scale with transformation (Eq. 7.4.4): gravity. The hydrodynamic equations (Eq. 7.4.1) are unchanged by scale transformation (7.4.4) only if gravitational acceleration transforms as  $g'_i = \alpha^{-1} g_i$ . Thus, if the experiment with a 1-cm diameter projectile described above is performed in a 1-g gravitational field, the scaled-up results apply to a 50-km diameter projectile excavating a crater in a  $2 \times 10^{-7}$  g gravitational field. Such scaled laboratory experiments thus have only limited applicability to planetary impact processes.

One solution to this predicament is to raise the acceleration of gravity in the laboratory. This is commonly done with centrifuges that currently attain accelerations of about 500 g. Scaling factors  $\alpha$  between 500 (to earth surface gravity) and 3000 (to lunar surface gravity) are thus realizable. Since modern centrifuges can accommodate craters up to almost a meter in diameter, centrifuge techniques push direct laboratory modeling of cratering processes to kilometer-size planetary craters. Craters in the tens to hundreds of km diameter range must still be approached through other forms of scaling or theoretical modeling.

Lampson's empirical  $W^{1/3}$  scaling turns out to be a direct consequence of the scale transformation (Eq. 7.4.4). Thus, when all distance scales are increased by  $\alpha$ , the explosive charge mass (or, equivalently, energy release)  $W$  increases as  $\alpha^3$ . Both crater diameter  $D$  and  $W^{1/3}$  change by the same factor, establishing the proportionality (Eq. 7.2.1). A similar argument applies to impacts, where  $W = \frac{1}{2}mv^2$ : transformation (Eq. 7.4.4) leaves  $v$  unchanged, but increases the projectile mass  $m$  by a factor  $\alpha^3$ . The fact that large explosion craters are scaled more accurately by a  $1/3.4$  power than by  $1/3$  is because gravity destroys the hydrodynamic equations' invariance to scale transformations.

The effect of gravity on crater size can be roughly estimated from energy balance. The major energy expenditure in excavating a large crater is overcoming the gravitational potential energy of the material that occupies the crater-to-be. This energy is equal to excavated mass  $m_e$  times  $g$  times the vertical distance  $H$  over which it is raised, which is proportional to  $D$ . Or, since  $m_e$  is proportional to  $\rho D^3$ , and  $H$  is proportional to  $D$ , the energy is proportional to  $\rho g D^4$ . Since this gravitational energy of excavation increases as  $D^4$  with increasing crater diameter, whereas other energy sinks such as the conversion of kinetic energy into heat increase only as  $D^3$ , the  $D^4$  term eventually overwhelms all others to dominate the energy budget. The excavation energy thus approaches some constant fraction  $f$  of the total impact energy  $W$  at large diameters, so that

$$D = (fW/\rho g)^{1/4} \quad (7.4.5)$$

The  $1/3.4$  power-law exponent derived from nuclear explosion craters is intermediate between the small crater exponent of  $1/3$  and the large crater exponent of  $1/4$ .

Recent experimental work, however, has shown that impact crater excavation depends upon momentum as well as energy, so this large-size energy scaling limit is probably never realized. The dimensions of impact craters in water, whose excavation is entirely controlled by gravitational potential energy, scale as  $W^{1/3.83}$ , not  $W^{1/4}$ .

### 7.5 DIMENSIONLESS RATIOS

Dimensionless ratios are commonly used in hydrodynamics to estimate the relative sizes of terms in the hydrodynamic equations and thus evaluate the importance of different physical processes. They are similarly useful in the study of impact cratering, because they indicate flow regimes dominated by different processes. The oldest and most familiar dimensionless ratio is the Reynolds number,  $Re$ , which is the ratio between inertial stresses, of order  $\rho v^2$ , and viscous stresses, of order  $\eta v/l$ , where  $\eta$  is viscosity and  $l$  is a characteristic length. Viscous stresses, if present, are contained in the deviatoric stress tensor  $\Sigma_{ij}$  in Equations 7.4.1a–7.4.1c. The Reynolds number is thus

$$Re = \frac{\rho v l}{\eta} \quad (7.5.1)$$

Viscosity is seldom important in the early stages of impact cratering, except for microscopic craters. Consider a 10 km/second impact in which silicate melt is produced with a viscosity of  $10^3$  Pa-second. The Reynolds number approaches 1 only for craters with characteristic dimensions  $l \lesssim 3 \times 10^{-5}$  m ( $30 \mu\text{m}$ ). Melt viscosity effects are important in the excavation of the microcraters observed on lunar surface rocks (see Sec. 2.2). Planetary scale craters are affected by the viscosity of target materials only in the last stages of excavation when flow velocities are small.

A more important ratio for impact cratering is the ratio  $R_s$  between inertial stress and material strength,  $Y$ . Strength is also contained in the stress tensor  $\Sigma_{ij}$ :

$$R_s = \frac{\rho v^2}{Y} \quad (7.5.2)$$

Material strength effects have a significant effect on the cratering flow when  $R_s \approx 1$ . In rock with a crushing strength of, for example, 100 MPa, material strength begins to influence the excavation flow when its velocity falls to about 200 m/second.

The importance of gravity is gauged by the Froude number  $F$ , which is the ratio of inertial and gravitational stresses,

$$F = \frac{v^2}{gl} \quad (7.5.3)$$

Thus, in a 100-km diameter crater on the moon, gravity modifies the cratering flow when  $F \lesssim 1$ , or for velocities less than about 400 m/second. Gravity becomes relatively more important as the size scale  $l$  increases, other conditions being the same.

Many other dimensionless ratios can be constructed from the quantities in the hydrodynamic equations. The last ratio of importance in impact

catering, however, is the ratio between gravitational stresses,  $\rho gl$ , and strength,  $Y$ . Denoting this ratio  $R_G$ ,

$$R_G = \frac{\rho gl}{Y} \quad (7.5.4)$$

This ratio marks a boundary between two different cratering regimes. When  $R_G < 1$ , strength dominates crater excavation. When  $R_G > 1$ , gravity dominates excavation. Scaling relations are expected to differ in these two regimes because the physical processes halting crater growth are different. Taking  $l \approx D$  and  $Y \approx 2$  MPa, equal to the observed yield stress at the onset of crater collapse in the fractured rock surrounding a freshly formed crater, the transition between strength and gravity-dominated regimes should occur at about 70 m on the earth and 400 m on the moon. Even if the yield strength  $Y$  is taken to be 100 MPa, equal to the crushing strength of intact rock, these numbers change to 3 km for the earth and 20 km for the moon, still smaller than many large impact craters. The growth of most planetary craters is thus dominated by gravity.

### 7.5 LATE-STAGE EQUIVALENCE

Late-stage equivalence was discovered in the late 1960s by J. K. Dienes and J. M. Walsh during the course of two-dimensional numerical code calculations. These computations simulated vertical impacts on strength-dominated targets, so their relevance to more realistic oblique impacts on gravity-dominated planets is questionable. However, the scaling implied by late-stage equivalence is useful for smaller, strength-dominated craters. Furthermore, late-stage equivalence forms the foundation on which the modern coupling constant approach to scaling is built (see below).

The numerical code calculations showed that the flow field late in the calculation, after the initial shock wave had dissipated and the cratering flow was well established, is unchanged if the projectile diameter  $L$  and impact velocity  $v_i$  are varied in such a way that

$$Lv_i^n = \text{constant} \quad (7.6.1)$$

where  $n = 0.58 \pm 0.01$ . This scaling law goes beyond the hydrodynamic similarity transformation (Eq. 7.4.4) because it relates impacts at different velocities to one another. Hydrodynamic similarity only relates different size impacts at the same velocity. Equation 7.6.1 is valid for hypervelocity impacts, when  $v_i \gtrsim 3.5$  km/second.

Late-stage equivalence leads to a scaling equation for crater diameter  $D$ , depth  $H$ , some other characteristic distance,

$$D/L = K(\rho_p/\rho_i)^{1/3}(v_i/c)^n \quad (7.6.2)$$

where  $c$  is the speed of sound in the target and  $K$  is a dimensionless function of material strength.

Although  $K$  is constant for a given target material, it must be determined by experiment. Each quantity of interest,  $D$ ,  $H$ , etc., has its own function  $K$ . Late-stage equivalence is probably violated when gravity becomes important in the flow near the end of crater excavation: it holds at all times only in the strength regime.

The "lateness" of Late-stage equivalence has been reinterpreted in modern work as an "intermediate" stage of cratering flow. As illustrated in Section 5.5.3, the crater grows as a simple function of time after the end of the contact and compression stage and before gravity or strength brings the cratering flow to a halt. Late-stage equivalence seems to hold universally in this regime and forms the basis for modern work on scaling relations via the "coupling constant" approach described below.

A scaling relation similar to that of Equation 7.6.2 was discovered experimentally by A. C. Charters and J. L. Summers (1959), who studied impacts into ductile metal targets at velocities between about 1 and 4 km/second. Charters and Summers found that the depth  $H$  of the final crater in pure metal targets (lead, aluminum, tin, zinc, copper) is given by a relation.

$$H/L = 2.28 \left( \frac{\rho_p}{\rho_t} \right)^{2/3} \left( \frac{v_i}{c} \right)^{2/3} \quad (7.6.3)$$

where the experiments covered the range of  $v_i/c = 0.2$  to  $3.0$ . The craters formed in the metal targets were generally hemispherical for projectiles of the same material as the target, deeper than hemispheres for projectiles denser than the target, and shallower than hemispheres for projectiles less dense than the target. Since the crater shape in these experiments thus depends somewhat upon the projectile's properties, the hypervelocity impact regime where late-stage equivalence is supposed to hold was evidently not quite reached. Nevertheless, the velocity exponent in Equation 7.6.3,  $0.66$ , is reasonably close to the late-stage equivalence value  $n = 0.58$ . The difference in exponent of the  $(\rho_p/\rho_t)$  terms in Equations 7.6.2 and 7.6.3 may be attributed to the dependence of the crater shape on projectile density.

Although Charters and Summers found that Equation 7.6.3 fit their data well for pure metals, there are large deviations when the target is a metal alloy such as hardened aluminum, steel, or an Mg-Li alloy. They attributed this to the inappropriate choice of  $c$ , the sound velocity in the target, to normalize the impact velocity  $v_i$ . Charters and Summers reasoned that the crater growth was eventually stopped by the target's plastic yield strength  $Y$ , which should thus appear in the scaling equation instead of the sound velocity  $c$ . They pointed out that the sound velocity is directly related to strength only for pure metals and that in alloys the ratio between strength and sound velocity may be far higher than in pure

metals. Their results for both alloys and pure metals do in fact fit a power scaling law much better if  $c$  in Equation 7.6.3 is replaced by  $(Y/\rho_t)^{1/2}$ .

### 7.7 PI-GROUP SCALING

The most recent approach to impact crater scaling has used the "Pi-theorem" originally devised by E. Buckingham in 1914 and elaborated by P. W. Bridgeman in 1949. This approach is essentially a sophisticated form of dimensional analysis in which the physically relevant parameters in an impact cratering event, such as transient crater diameter  $D_{at}$  or volume  $V$ , impact velocity  $v_i$ , target and projectile densities  $\rho_p$  and  $\rho_t$ , target strength  $Y$  and gravity  $g$ , and projectile mass  $m$  are combined into a number of dimensionless parameters. The functional dependence of one parameter on the others may then be determined experimentally, holding all but the parameters of interest fixed.

For example, the apparent diameter  $D_{at}$  of a transient crater can be expressed as a function of all the variables above:

$$D_{at} = F(v_i, \rho_p, \rho_t, Y, g, m) \quad (7.7.1)$$

The crater diameter  $D_{at}$  appears to be a function of five independent variables. However, because all physical quantities can be expressed in terms of mass, length, and time, and because the function  $F$  must have the same dimension as the left-hand side,  $D_{at}$  is actually a function of only three dimensionless ratios.

Choose dimensionless combinations

$$\pi_D = D_{at} \left( \frac{\rho_t}{m} \right)^{1/3} \quad (7.7.2a)$$

$$\pi_2 = \frac{1.61 g L}{v_i^2} \quad (7.7.2b)$$

$$\pi_3 = \frac{Y}{\rho_p v_i^2} \quad (7.7.2c)$$

$$\pi_4 = \rho_t / \rho_p \quad (7.7.2d)$$

where  $\pi_D$  is a dimensionless measure of the crater diameter, roughly equal to the ratio of crater diameter to projectile diameter ( $\pi_D = 1.24 (\rho_t / \rho_p)^{1/3} D/L$  for a spherical projectile),  $\pi_2$ , the "gravity-scaled size" is essentially the inverse of the Froude number  $F$  (Eq. 7.5.3),  $\pi_3$  is the inverse of  $R_s$  (Eq. 7.5.2) and measures the influence of target strength, and  $\pi_4$  is the ratio of target and projectile densities.

The functional relation (Eq. 7.7.1) can now be written in its most general dimensionless form:

$$\pi_D = F'(\pi_2, \pi_3, \pi_4) \quad (7.7.3)$$

Two impacts are "similar" if they have the same values of dimensionless parameters in the argument of Equation 7.7.3, even though the

actual impact events may differ in velocity, size, gravitational acceleration or target strength. The validity of the nondimensionalization procedure is tested experimentally by verifying that "similar" sets of parameters result in craters with the same scaled diameter. Such experiments have met with considerable success.

This relation becomes still simpler if craters are in the gravity regime where, by definition,  $\pi_3$  is negligible, and if the ratio of projectile and target densities  $\pi_4$  is constant. In this case

$$\pi_D = F''(\pi_2) \quad (7.7.7)$$

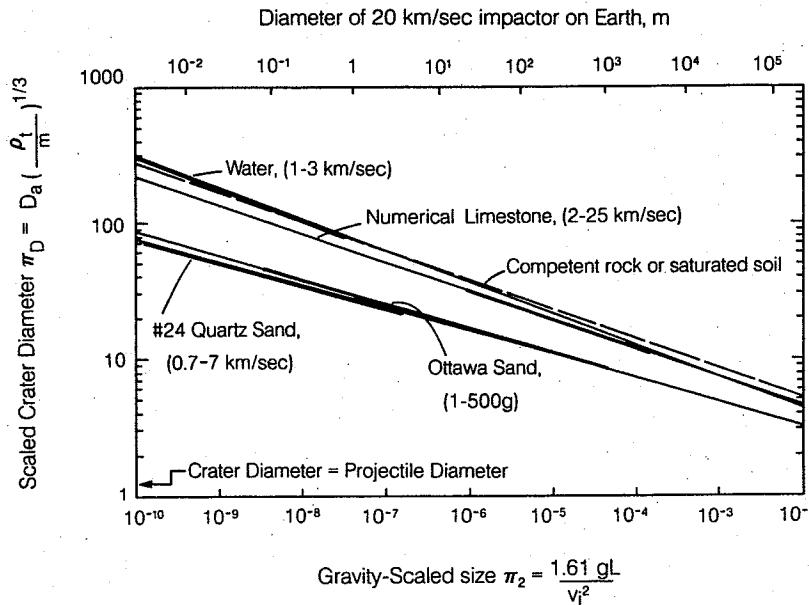
The simple equation (7.7.7), which involves only two dimensionless variables  $\pi_D$  and  $\pi_2$ , actually incorporates six dimensionless quantities:  $D$ ,  $\rho_p$ ,  $m$ ,  $g$ ,  $v_i$ , and  $L$ . Dimensional analysis and elementary physical insight have greatly reduced the complexity of the equation.

The validity of Equation 7.7.7 has been assessed experimentally. Figure 7.3 plots  $\pi_D$  against  $\pi_2$  for a large number of impact cratering data from a variety of sources, ranging from the impact of pyrex spheres into water at 1 to 3 km/second to impact experiments performed at 500 g in sand. The figure includes data obtained at less than 1 g acceleration in drop buckets as well as the results of a series of hydrocode calculations. It is

clear that the data form a number of discrete linear arrays on these log-log plots.

The fact that the data do not follow a single functional trend means that some dimensionless factor has been missed in Equation 7.7.7. Since data for a single material do follow a single line, the missing factor must be a material parameter. This parameter was originally believed to be the angle of internal friction of the target material. As a result of further experiments and theoretical studies, however, the principal missing factor is now thought to be the target's porosity,  $\phi$ , which may vary considerably among loose granular materials such as sand. Internal friction may also play a role: present experiments are not able to cleanly separate the effects of internal friction and porosity. A new dimensionless parameter could be defined, say,  $\pi_5 = \phi$  that expresses this dependence. Granular materials with the same porosity and internal friction should define a single curve on this plot, as observed.

All of the data plotted in Figure 7.3 are for vertical impacts. The projectile's angle of incidence  $\theta$  is another dimensionless parameter that must be included in a full description of impact cratering. Although this is easy to do mathematically, few experiments have yet been performed for oblique impacts.



**Fig. 7.3** Impact crater scaled diameter  $\pi_D$  vs. gravity-scaled size,  $\pi_2$ . The heavy portions of the lines are determined by experimental data and the light portions are extrapolations. The heavy line width is larger than the dispersion of the data. Data included is from impacts in water at 1 to 3 km/second, impacts into two different types of sand under both microgravity and 500 g centrifuge conditions, and the results of numerical computations. The data for any one material follows a straight line on these log-log plots, thus supporting a power-law relationship between  $\pi_D$  and  $\pi_2$ . The near-identity of the lines for competent rock and water is deceptive: although the *diameter* of craters in water and competent rock is similar, water craters are hemispherical, so the *volume* of craters in water is about three times larger than craters in rock. The upper scale shows the projectile diameter for a 20 km/second impact on earth. After Schmidt (1980) and Schmidt and Housen (1987).

The straight-line dependence of the data in Figure 7.3 indicates that the function (Eq. 7.7.7) is a simple power law of form:

$$\pi_D = C_D \pi_2^{-\beta} \quad (7.7.8)$$

Where  $C_D$  and  $\beta$  are experimentally determined constants.

Experimenters generally prefer to quote their cratering results in terms of volume excavated, not crater diameter. Diameter is emphasized in this book because it is more easily determined from photogeologic data than crater volume and is less likely to be altered by erosion or lava infilling of the crater. However, since most experimental cratering data are reported in terms of volume or mass displaced, it is useful to define the dimensionless parameter  $\pi_V$ :

$$\pi_V = \rho_i V/m \quad (7.7.9)$$

This quantity is the ratio between the mass displaced from the crater and the mass of the projectile. It is commonly called the "cratering efficiency." Cratering efficiency is also a power function of  $\pi_2$  in the gravity regime:

$$\pi_V = C_V \pi_2^{-\gamma} \quad (7.7.10)$$

Table 7.1 summarizes the latest experimental determinations of the constants in Equations 7.7.8 and 7.7.10 for a variety of target materials.

Impact crater diameter and volume are related by simple geometric factors. For example, the volume  $V$  of a hemispherical crater is related to its diameter  $D$  by  $V = \pi D^3/12$ . Most transient craters in earth materials approach a parabolic shape for which transient crater depth  $H_{at}$  is roughly one-third to one-fourth the diameter  $D_{at}$ . The volume of a parabolic crater is  $V = \pi H_{at} D_{at}^2/8$ . If  $H_{at}/D_{at}$  is independent of crater size,  $\pi_V$  and  $\pi_D$  are related by:

$$\pi_D = \left( \frac{\pi H_{at}}{8D_{at}} \right)^{1/3} \pi_V^{1/3} \quad (7.7.11)$$

TABLE 7.1 *Experimental power law scaling constants\**

Target Material	Diameter Constants†		Volume Constants‡	
	$C_D$	$\beta$	$C_V$	$\gamma$
Water	1.88	0.22	2.1	0.65
Quartz Sand	1.4	0.16	0.4	0.49
Ottawa Sand	1.68	0.17	0.24	0.51
Competent Rock or Saturated Soil	1.6	0.22	0.2	0.65

†Equation 7.7.8 of the text

‡Equation 7.7.10 of the text

\*From Schmidt (1980) and Schmidt and Housen (1987).

The data in Table 6.1 indicates that the exponent  $\beta$  is equal to  $\gamma/3$  within experimental uncertainty, which supports the constancy of  $H_{at}/D_{at}$ .

The observed dependence of  $\pi_D$  or  $\pi_V$  on a simple power of  $\pi_2$  is not explained by Pi-group scaling or simple invariances of the hydrodynamic equations. The validity of power laws appears to be due to the existence of a single "coupling parameter" that links the early- and late-time cratering flow fields. The "early-time" flow encompasses the contact and compression stage and part of the excavation stage, as defined in Chapters 4 and 5. During this part of the cratering event, details of the projectile shape, mass, and impact velocity are gradually lost as the flow develops. Late-time parameters such as crater diameter, depth, and gravitational acceleration are affected only by a dimensional combination of these factors, the "coupling parameter," which, up to a power of projectile density, turns out to be  $Lv_i^n$ , the same parameter defined by late-stage equivalence. In this context "Late Stage" is really intermediate in the cratering flow, after contact and compression and before the cessation of motion around the crater. If there is only one such dimensional combination that couples the early-time and late-time flows, it can be shown (Holsapple and Schmidt, 1987) that scaling relations among late-time variables must take the form of power laws. Further work is needed before the precise significance of the coupling parameter is clarified. However, the approximate power-law dependence of crater scaling laws is undeniable. The coupling constant approach may explain the prevalence of such relations in impact cratering. These power laws may fail only for the very largest craters where gravity is so important that crater excavation takes place entirely during the contact and compression stage, before a coupling parameter is defined.

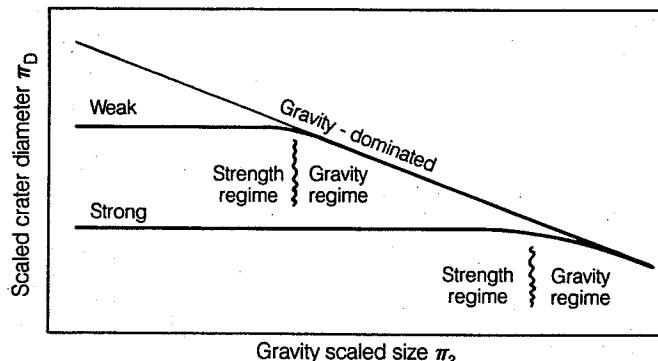
The diameters or volumes of small craters in hard rock or other materials with significant strength do not scale as a function of the  $\pi_2$  parameter. Instead, crater dimensions in the "strength regime" are a function of the  $\pi_3$  parameter (Eq. 7.7.2c). Observations, and the coupling constant approach, show that the scaled crater diameter is a power function of  $\pi_3$ ,

$$\pi_D = C'_D \pi_3^{-\sigma} \quad (7.7.12)$$

where the coupling constant theory predicts that the exponent  $\sigma$  is related to the exponent  $\beta$  of gravity regime craters (Eq. 7.7.8 and Table 7.1) by

$$\sigma = \beta/(\beta - 1) \quad (7.7.13)$$

As crater size increases, other things being the same, gravity becomes relatively more important in crater excavation and the crater diameter is eventually controlled by gravity alone. Figure 7.4 shows how this transition appears on a plot of scaled-crater diameter against gravity-scaled size. Below the transition, crater diameter is indepen-



**Fig. 7.4** Transition from the strength to gravity scaling regimes. The scaled crater diameter  $\pi_D$  is independent of the gravity-scaled projectile size  $\pi_2$  in the strength regime, but as  $\pi_2$  increases gravity eventually dominates. For similar projectile size and velocity the transient crater diameter is larger in weak materials than in strong ones.

dent of  $\pi_2$  and the scaled crater diameter plots as a horizontal line. The weaker the material in the strength regime, the larger the crater. The crater diameter at the transition between the strength and gravity regimes is independent of velocity (this is apparent even from the relatively simple considerations of Equation 7.5.4). It can be shown (Chapman and McKinnon, 1986) that this transition diameter  $D_{TR}$  scales as

$$D_{TR} = C_{TR} Y g^{-1} \rho_p^{-2/3} \rho_t^{-1/3} \quad (7.7.14)$$

where  $C_{TR}$  is a constant to be determined by experiment. Unfortunately, it is not yet well known. Experiments on the strength-gravity transition in basalt targets have shown that the combination  $C_{TR} Y \approx 0.9$  MPa. Taking  $Y = 100$  MPa for basalt implies  $C_{TR} \approx 0.009$ . Unfortunately, the relevant "yield strength"  $Y$  for basalt is not well known, so only the combination  $C_{TR} Y$  is meaningful. The experimental value of this combination implies that the strength-gravity transition occurs for craters about 220 m in diameter on the lunar mare, assuming them to be formed by carbonaceous chondrite ( $\rho_p = 2200$  kg/m<sup>3</sup>) projectiles.

## 7.8 COMPARISON OF SCALING RELATIONS FOR VERTICAL IMPACTS

So far, two different approaches to crater diameter scaling when gravity is important have been discussed: Lampson's yield scaling and Holsapple and Schmidt's (1982) Pi-group scaling. Another set of scaling relations based on energy expended in cratering was developed by D. E. Gault. These relations summarize his and his co-workers results at the NASA Ames Vertical Gun Facility up to 1974. This section compares yield scaling, Gault's scaling laws, and Pi-group scaling for similar impact conditions over a range of crater sizes.

Gault's (1974) summary of his scaling laws applies in three different ranges of crater size and

materials. For craters up to about 10 m in diameter in solid rock, he finds:

$$D_{at} = 0.015 \rho_p^{1/6} \rho_t^{-1/2} W^{0.37} (\sin\theta)^{2/3} \quad (7.8.1a)$$

For craters up to roughly 100 m in diameter in loose soil or regolith,

$$D_{at} = 0.25 \rho_p^{1/6} \rho_t^{-1/2} W^{0.29} (\sin\theta)^{1/3} \quad (7.8.1b)$$

And for all craters larger than 100 m to 1 km, in any type of target material,

$$D_{at} = 0.27 \rho_p^{1/6} \rho_t^{-1/2} W^{0.28} (\sin\theta)^{1/3} \quad (7.8.1c)$$

These relations are supposed to grade smoothly into one another over some range of transitional diameters. All units are in the mks system. Gravity does not explicitly enter these equations: the above formulae are valid for the moon, with  $g = 1.67$  m/second<sup>2</sup>. The equations may be scaled to other bodies by assuming that  $D_{at}$  is proportional to  $g^{-0.165}$ , a dependence that Gault found in other experiments.

The modern form of yield scaling derives from nuclear explosion data. Nordyke (1962) gives the following relation between scaled explosion crater diameter  $D_{at}/W^{1/3.4}$  and scaled depth of burial  $Z_s = d/W^{1/3.4}$ :

$$D_{at}/W^{1/3.4} = 0.0133 + 1.51Z_s - 5.5 \times 10^3 Z_s^3 - 1.49 \times 10^6 Z_s^5 \quad (7.8.2)$$

where units are in the mks system. This equation holds for depths of burial from zero (surface burst) to the depth for which the crater diameter is zero. Impact craters typically have a very shallow equivalent depth of burial so that only the first term in  $Z_s$  needs to be evaluated. Following Equation 7.3.1, set  $Z_s \approx \rho_p^{1/2} \rho_t^{-1/2} L / W^{1/3.4}$  to obtain an equation for impact-crater diameter from yield scaling:

$$D_{at} = 0.0133 W^{1/3.4} + 1.51 \rho_p^{1/2} \rho_t^{-1/2} L \quad (7.8.3)$$

TABLE 7.2 Comparison of three impact crater scaling laws for a 20 km/sec impact on earth

Projectile Diameter	Transient Crater Diameter (km)		
	Yield Scaling* (1962)	Gault† (1974)	Schmidt‡ Table 7.1 (1987)
1 m	0.034	0.028	N/A
10 m	0.26	0.21	0.38
100 m	2.0	1.1	2.3
1 km	16.0	7.6	14.0
10 km	120.0	53.0	84.0
100 km	975.0	364.0	510.0

\*Equation 7.8.3

†Equation 7.8.1a to 7.8.1c

‡Equation 7.8.4

Because this equation was derived from terrestrial explosion data, it corresponds to  $g = 9.8 \text{ m/second}^2$  and, naturally, has no dependence on the angle of impact  $\theta$ .

Finally, the Pi-group scaling relation for competent rock from Table 7.1 can be rearranged into the form:

$$D_{at} = 1.8 \rho_p^{0.11} \rho_i^{-1/3} g^{-0.22} L^{0.13} W^{0.22} \quad (7.8.4)$$

This equation contains an explicit  $g$  dependence but is valid only for vertical impacts,  $\theta = 90$  degrees.

Table 7.2 compares the predictions of the 1962 yield scaling law (Eq. 7.8.3), Gault's scaling law (Eq. 7.8.1), and the 1983 Schmidt-Holsapple Pi-group scaling law (Eq. 7.8.4) for comparable conditions in impacts on the earth. The largest differences among the three laws occur, not surprisingly, at the largest crater sizes. Even for projectiles 100 km in diameter, however, the maximum difference in predicted crater diameter is less than a factor of 3. The agreement is fairly good among these three scaling laws in spite of their great differences in origin and philosophy. Part of the reason for this agreement is that transient crater diameter is not a very sensitive function of impact energy: for all three laws  $W$  is raised to a fractional power less than one-third. Although this is good for estimates of crater diameter from impact energy, or energy and velocity together, it also means that the determination of an impact energy from a crater diameter is very imprecise. A factor 3 uncertainty in crater diameter, raised to the 3.4 power, is roughly a factor 40 uncertainty in energy. Much more confidence in scaling relations must thus be obtained before energy estimates can be considered firm. The history of Meteor Crater formation energy estimates in Section 7.3 is indicative of the state of flux in this area.

## 7.9 OBLIQUE IMPACTS

Although little theoretical work has been done on scaling oblique impacts, there is a body of experimental data on less-than-meter-size craters in sand and rock. This data was collected largely by D. E. Gault and J. A. Wedekind at NASA's Ames Research Laboratories. Additional data was collected by H. J. Moore (1976) from missile-impact craters ranging up to 12 m in diameter at White Sands Proving Ground in New Mexico.

Some of Gault and Wedekind's data is presented in Figure 7.5, where the ratio of the mass displaced from a crater in noncohesive quartz sand at impact angle  $\theta$  to the mass displaced by a vertical impact ( $\theta = 90$  degrees) is plotted as a function of impact angle. The impact angle ranges from 90 degrees down to 2 degrees. This data strongly supports a simple  $\sin\theta$  dependence of this ratio. Similar data for solid rock, where target strength dominates crater formation, yields an approximate  $\sin^2\theta$  dependence.

Gault and Wedekind's findings imply that the cratering efficiency  $\pi_v$  is directly proportional to  $\sin\theta$ . The dimensionless crater diameter  $\pi_D$  is therefore proportional to  $(\sin \theta)^{1/3}$  by Equation 7.7.11, so long as the depth/diameter ratio  $H_{at}/D_{at}$  of oblique impact craters is the same as for vertical impacts. This dependence of diameter on impact angle is explicit in Gault's scaling law, Equations 7.8.1a-7.8.1c.

A number of workers have suggested that only the vertical component of the impact velocity,  $v_i \sin \theta$ , is effective in excavating the crater. In this case  $\pi_v$  is proportional to  $(\sin \theta)^2$  and  $\pi_D$  is proportional to  $(\sin \theta)^{2/3}$ . The power-law data in Table 7.1 shows that this is essentially the same as Gault's scaling law for sand targets, which Gault

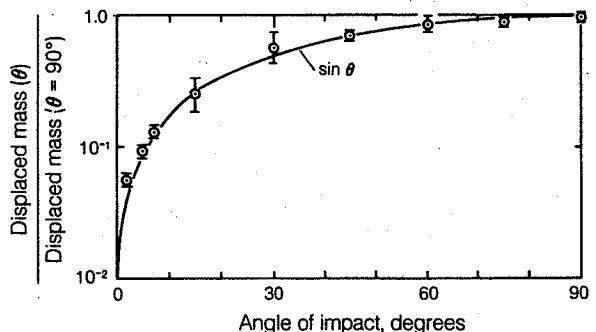


Fig. 7.5 The ratio between mass displaced from an oblique impact crater to the mass displaced from a vertical impact is a function of impact angle  $\theta$ . This data was derived from 444 rounds of either aluminum or pyrex spheres impacting non-cohesive quartz sand targets at velocities from 1.8 to 6.8 km/second. This data supports a  $\sin\theta$  dependence of  $\pi_v$ . After Gault and Wedekind (1978).

used in his experiments, but that there are substantial differences for water and large craters in competent rock or saturated soils. At the moment no one can say which form is more correct.

### 7.10 MELT AND VAPOR VOLUME

The quantity of melt or vapor produced in an impact has not yet been determined experimentally because the threshold impact velocity for significant melting or vaporization cannot yet be attained in the laboratory. Theoretical computations described in Section 5.2.3 indicate that once the production threshold has been reached, the ratio between the mass of melt or vapor and the projectile mass scales as the square of the impact velocity,  $v_i^2$ . The effect of obliquity on the production of melt or vapor is poorly known, although jetting in oblique impacts may actually *enhance* the mass produced.

O'Keefe and Ahrens (1982b) present computations of melt and vapor masses for the impact of iron, gabbroic anorthosite, and water projectiles on gabbroic anorthosite (a rock simulating the lunar crust) targets. Their results can be summarized in two equations that are valid when the predicted mass of melt or vapor is several times larger than the projectile mass.

$$\frac{\text{Mass of Melt}}{\text{Mass of Proj.}}$$

$$= 0.14 \frac{v_i^2}{\epsilon_m} \text{ for } v_i \gtrsim 12 \text{ km/second} \quad (7.10.1a)$$

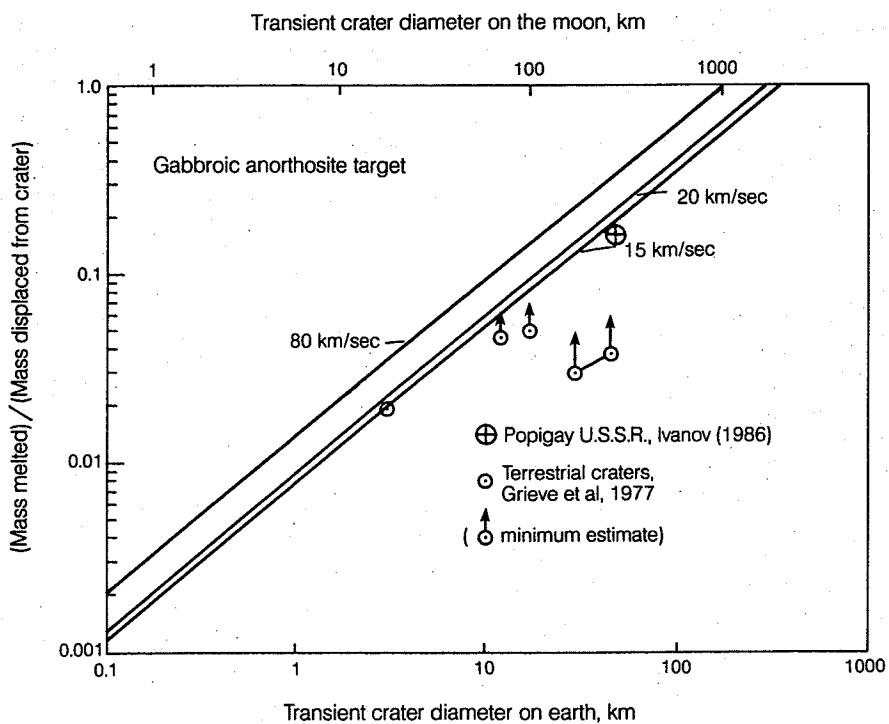
and

$$\frac{\text{Mass of Vapor}}{\text{Mass of Proj.}}$$

$$= 0.4 \frac{v_i^2}{\epsilon_v} \text{ for } v_i \gtrsim 35 \text{ km/second} \quad (7.10.1b)$$

where  $\epsilon_m$  and  $\epsilon_v$  are the specific internal energies of melting and vaporization of the target. For gabbroic anorthosite  $\epsilon_m = 3.4 \times 10^6 \text{ J/Kg}$  and  $\epsilon_v = 5.7 \times 10^7 \text{ J/Kg}$ . Although the coefficient in Equation 7.10.1a is smaller than the coefficient in Equation 7.10.1b, the mass of melt produced in an impact is always larger than the vapor mass by nearly a factor of 10. The deceptive difference in the size of the coefficients is more than compensated by the different values of  $\epsilon_m$  and  $\epsilon_v$  in the denominators. Impacts onto targets more volatile than gabbroic anorthosite presumably vaporize more material, although this has not yet been verified by computations.

The mass of material melted or vaporized is de-



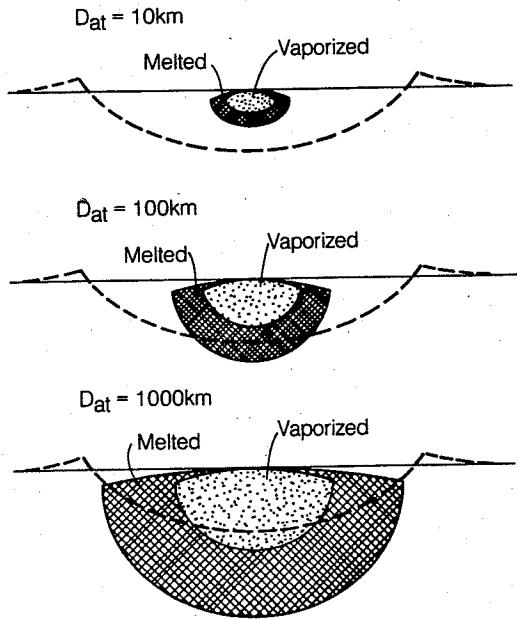
**Fig 7.6** The ratio between the mass melted in an impact into gabbroic anorthosite and the mass displaced from a crater is plotted against crater diameter on the earth (lower horizontal scale) and the moon (upper horizontal scale). The mass of melt increases with crater size and impact velocity. The points show melt volumes estimated for four terrestrial craters by Grieve et al. (1977). Except for Brent Crater, these are minimum estimates. They appear to be consistent with the theoretical curves, which are plotted from Equation 7.10.2 in the text.

terminated in the early high-pressure phase of shock wave expansion. This mass is independent of either material strength or gravity. The final size of large craters, however, is determined by gravity. Since the cratering efficiency  $\pi_v$  decreases with increasing crater size, the mass of melt of vapor makes up a progressively larger proportion of the total mass displaced as crater size increases. Sufficiently large craters may form entirely within the mass of shock melted material. Such craters would have a very different morphology than smaller craters: for example, a conventional rim would not form, and nearly all of the ejecta would be molten.

Equations 7.10.1, 7.7.10, and 7.7.8 may be combined with the scaling constants in Table 7.1 for competent rock to derive a general formula for the ratio between the mass of rock melted to the mass displaced from the crater (SI units);

$$\frac{\text{Mass Melted}}{\text{Mass Displaced}} = 1.6 \times 10^{-7} (gD_{at})^{0.83} v_i^{0.33} \quad (7.10.2)$$

This equation applies only to the target material gabbroic anorthosite. The melt volume for water ice targets, for example, is probably larger because of water's smaller energy of melting  $\epsilon_m$  (estimated at  $2 \times 10^6 \text{ J/Kg}$ ).



**Fig. 7.7** Scaling laws 7.10.1 and 7.10.2 imply that as crater diameter increases the volume of melted and vaporized material may eventually approach the volume of the crater itself. Although melt and vapor may be only a minor component of the ejecta from small craters, extremely large craters may form almost entirely in a melt pool. This figure is constructed for impacts at 35 km/second on the earth.

The ratio between the mass of target material melted and the mass displaced from the crater is plotted in Figure 7.6 as a function of crater diameter for several impact velocities on the earth and moon. Larger craters clearly have a larger proportion of melt, illustrated for terrestrial craters in Fig. 7.7. The same statement holds true for vapor when the impact velocity is large enough for vaporization to be important. The melt mass is equal to the mass displaced for craters more than about 300 km in diameter on the earth for an impact velocity of 20 km/second. Higher impact velocities naturally produce more melt and lower this diameter. On the moon the melt mass equals the mass displaced only for craters larger than 1800 km in diameter for impacts at 20 km/second.

### 7.11 CRATER FORMATION TIME

How long does a large crater take to form? Until recently there was no way to answer this question, since crater formation times were not usually measured and it was not known how to extrapolate laboratory measurements to large crater sizes. Recent work has improved the situation, however, and it is now possible to scale crater formation times within the context of Pi-groups.

A nondimensional crater formation time  $\pi_T$  is defined by the ratio of the formation time  $T_f$  to the time interval required for the projectile to traverse its own radius,  $L/2v_i$ :

$$\pi_T = 2T_f v_i / L \quad (7.11.1)$$

Laboratory experiments (Schmidt and Housen, 1987) support a power-law relationship between  $\pi_T$  and  $\pi_2$ ,

$$\pi_T = C_T \pi_2^\tau \quad (7.11.2)$$

where the constants  $C_T$  and  $\tau$  are measured to be 2.3 and 0.61, respectively, for water and 1.6 and 0.61, respectively, for wet sand. Both types of crater are affected only by gravity; strength plays no role. The similarity of the slopes  $\tau$  for these two very different target materials may be due to the existence of a single coupling parameter. The coupling-parameter approach predicts that  $T_f$  is proportional to  $(D_{at}/g)^{1/2}$ , which is also roughly the period of a wave of wavelength  $D_{at}$  in a deep inviscid fluid, a physically plausible result. The numerical factor in this expression is  $T_f \approx 0.54 (D_{at}/g)^{1/2}$  for these gravity-dominated craters.

The formation time of a 2.3-km diameter crater on earth produced by the impact of a 100-m diameter projectile at 20 km/second is, by Equation 7.11.2, roughly 10 seconds. This is considerably longer than the 0.005-second duration of the contact and compression stage. An 84-km diameter crater produced by a 10-km diameter projectile striking at 20 km/second forms in about 64 seconds.

The formation time  $T_f$  is the time required for the *transient* crater to form: it does not include a possible subsequent collapse, although the collapse time itself is expected to scale as  $(D_{\text{av}}/g)^{1/2}$ .

Recent studies of the rate of crater growth, described in Section 5.5.3, show that, prior to the onset of strength or gravity effects, the crater depth and diameter increase as a power of time,  $H_{\text{at}}(t) \propto t^{0.4}$ . The coupling-parameter approach predicts that this power is given by  $n/(n + 1)$ , where  $n$  is the power of impact velocity in the late-stage equivalence parameter (Eq. 7.6.1) or, equivalently, the power is given by  $2\gamma/(3 + \gamma)$ , where  $\gamma$  is the power in the volume-scaling relation (Eq. 7.7.10). In either case, for  $n = 0.58$ , or  $\gamma = 0.65$ , the crater dimensions are predicted to grow as  $t^{0.36}$ , in good agreement with observations.

## 7.12 EJECTA SCALING

Experts on explosion cratering have long known that the debris blanket thrown from an explosion crater does not scale in a self-similar way with crater size. This conclusion follows directly from hydrodynamic similarity (Sec. 7.4), because the range of a fragment ejected with velocity  $v_e$  depends upon gravity, whereas the size of most explosion craters does not depend upon gravity (such craters are usually dominated by target strength). Thus a fragment thrown out at velocity  $v_e$  from, say, a position in the crater halfway between the center and the final rim travels the same distance from its starting position in both small and large craters. However, this distance is a different fraction of the crater diameter for small and large craters, so that the resulting ejecta deposit cannot be self-similar to the crater.

Lunar crater mappers, however, have long assumed that the ejecta deposits of lunar craters *do* scale in a self-similar fashion. This assumption is supported by a large volume of data that indicates that the radius of the outer edge of the "continuous" ejecta deposit, that is, the deposit that blankets and erases or obscures most preexisting topography, is close to 2.35 times the crater radius in craters ranging from 1.3 km to 436 km in diameter. The radius of the continuous ejecta departs from this trend, becoming nearly constant, only in craters smaller than 1.3 km in diameter.

This apparent self-similarity of ejecta blanket diameters around gravity-dominated craters appears to hold on all the planets and satellites studied so far. Although Mercurian ejecta blankets may be slightly less extensive than those around lunar craters, the ejecta blanket diameter is roughly equal to one crater diameter on Mars (except for craters with fluidized ejecta), the Jovian satellites, and even the tiny Saturnian satellites Dione and Rhea (see Section 6.3.1).

The apparent discrepancy between explosion craters and lunar craters is resolved by the recog-

nition that the explosion craters used to establish nonsimilarity are dominated by strength, whereas large lunar craters are dominated by gravity. Pi-group analysis shows that the ejection velocity  $v_e$  from a position  $r$  within a crater of final radius  $R$  is given by

$$\frac{v_e}{\sqrt{gR}} = F\left(\frac{r}{R}, \frac{Y}{\rho g R}\right) \quad (7.12.1)$$

This relation holds only after the excavation flow is well established and the effects of projectile-specific variables can be neglected. In particular, it is valid only for radius  $r$  larger than a few times the projectile radius  $L/2$ .

When the strength  $Y$  can be neglected, the function  $F$  in Equation 7.12.1 depends upon  $r/R$  alone. Furthermore, both the coupling parameter approach and the Z-model (Sec. 5.4.3) suggest that  $F(r/R)$  is a power of  $r/R$ :

$$\frac{v_e}{\sqrt{gR}} = C_e \left(\frac{r}{R}\right)^{-\epsilon} \quad (7.12.2)$$

The power  $\epsilon$  can be determined either experimentally or derived from the coupling-parameter model, where it is equal to  $(3 - \gamma)/2\gamma$ . In the later case it ranges from 1.9 for water to 2.4 for sand, exponents that are in reasonable agreement with Z-model predictions and with the limited experimental data that exists.

Equation 7.12.2 shows that the combination  $v_e/\sqrt{gR}$  is constant at a given value of  $r/R$ , however large or small the crater may be, so long as its excavation is dominated by gravity. Since the ballistic range of ejecta  $R_b = v_e^2 \sin 2\Phi/g$ , where  $\Phi$  is the ejection angle, the ratio of ejecta range to the crater radius  $R_b/R$  depends only upon  $r/R$ . That is, the ejecta of gravity-dominated craters scale self-similarly to the crater radius (or diameter).

Self-similarity means that debris thrown out, say, from a position halfway between the crater's center and rim land a fixed number of crater radii beyond the rim and there pile up to a height equal to a fixed fraction of the crater radius, whatever the size of the crater. This simple relation is violated when strength is important. It is also violated by the early high-velocity spall ejecta and jetted material, since both processes are governed by stress wave propagation and are therefore independent of gravity. It is noteworthy that the secondary crater fields of lunar and Mercurian craters are *not* self-similar (see Section 6.4), providing further evidence that secondary craters are produced by material ejected from the near-surface zone by stress wave "spall" interactions.

Although the thickness of the ejecta deposit may scale self-similarly, morphologic differences are expected as a function of crater size because the ejecta that land at a given scaled range from a large crater travel faster than those at the same

scaled range from a small crater. Thus, ejecta landing one crater diameter away from a 100-m diameter lunar crater strike at a velocity of only about 13 m/second, whereas the ejecta landing one crater diameter away from a 100-km crater strike at about 410 m/second. This difference in velocity affects the morphology of the deposit. High impact velocities, in particular, enhance secondary and tertiary cratering and may lead to the incorporation of large amounts of substrate material into the ejecta deposit (see Section 6.3.3).

Self-similarity supports the common contention that ejecta blanket thickness  $\delta$  falls off as a power of the distance  $r$  from the crater center. The usual relation  $\delta = f(R) (r/R)^{-3 \pm 1}$ , where  $f(R)$  is a function of crater radius only, is consistent with coupling-parameter predictions and brackets a large amount of experimental data.

A final quantity that appears to scale in a self-similar fashion for gravity-dominated craters is the volume of ejecta  $V_e$  with velocity equal to or greater than  $v_e$ . A power-law formula can be derived:

$$V_e/R^3 = C_{ej} \left( \frac{v_e}{\sqrt{gR}} \right)^v \quad (7.12.3)$$

where the power  $v$  can be shown to be  $6\gamma/(3 - \gamma)$  and ranges between about 1.7 for water and 1.2 for sand, again in good agreement with the limited experimental data that exists.

The field of ejecta scaling is relatively new and some changes may be expected in the scaling relations presented in this section as new data accumulates. No data yet exists for oblique impacts, for example. Nevertheless, many modern questions about impact cratering require detailed knowledge of crater ejecta and more data can be expected in the future.

## FURTHER READING

### General References

There is no comprehensive, up-to-date treatment of impact crater scaling in the literature today. The best treatment of modern research is found in

R. M. Schmidt, Meteor Crater: Energy of formation—Implications of centrifuge scaling, in *Proc. Lunar Planet. Sci. Conf. 11<sup>th</sup>*, pp. 2099–2128 (1980).

A fine discussion of the somewhat older energy-based scaling arguments is included in

D. E. Gault, J. E. Guest, J. B. Murray, D. Dzurisin, and M. C. Malin, Some compari-

sons of impact craters on Mercury and the Moon, *J. Geophys. Res.* 80: 2444–2460 (1975).

### Special Topics

A clear and accessible reference on yield scaling for explosions is

M. D. Nordyke, An analysis of cratering data from desert alluvium, *J. Geophys. Res.* 67: 1965–1974 (1962).

The impact-explosion analogy is addressed experimentally by

V. R. Oberbeck, Application of high explosion cratering data to planetary problems, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 45–65 (1977).

The equivalent depth of burial (sometimes called the “depth of burst”) relating explosions and impacts is analyzed by

K. A. Holsapple, The equivalent depth of burst for impact cratering, in *Proc. Lunar Planet. Sci. Conf. 11<sup>th</sup>*, pp. 2379–2401 (1980).

Late-stage equivalence is most thoroughly discussed in the original paper,

J. K. Dienes and J. M. Walsh, Theory of impact: Some general principles and the method of eulerian codes, in R. Kinslow (Ed.), *High-velocity impact phenomena*, Academic Press, New York pp. 45–104 (1970).

This paper also includes a discussion of hydrodynamic scaling, there called “dynamical scaling.”

The recent Pi-group formulation of impact-crater scaling is fully presented in

K. A. Holsapple and R. M. Schmidt, On the scaling of crater dimensions 2. impact processes, *J. Geophys. Res.* 87: 1849–1870 (1982).

Scaling of oblique impacts is one of the subjects of

D. E. Gault and J. A. Wedekind, Experimental studies of oblique impact, in *Proc. Lunar Planet. Sci. Conf. 9<sup>th</sup>*, pp. 3843–3875 (1978).

The relatively new subject of ejecta scaling was virtually created by

K. R. Housen, R. M. Schmidt, and K. A. Holsapple, Crater ejecta scaling laws: Fundamental forms based on dimensional analysis, *J. Geophys. Res.* 88: 2485–2499 (1983).

The new coupling-parameter approach to scaling, discussed briefly in the text, has finally appeared as a paper:

K. A. Holsapple and R. M. Schmidt, Point-source solutions and coupling parameters in cratering mechanics, *J. Geophys. Res.* 92: 6350–6376 (1987).

## VIII. CRATERING MECHANICS: MODIFICATION STAGE

### 8.1 INTRODUCTION

Shortly after the excavation flow has opened the transient crater and the ejecta has been launched onto ballistic trajectories, a major change takes place in the motion of debris within and beneath the crater. Instead of flowing upward and away from the crater center, the debris comes to a momentary halt, then begins to move downward and back toward the center whence it came. This collapse is generally attributed to gravity, although elastic rebound of the underlying, compressed rock layers may also play a role. The effects of collapse range from mere debris sliding and drain-back in small craters to wholesale alteration of the form of larger craters in which the floors rise, central peaks appear, and the rims sink down into wide zones of stepped terraces. Great mountain rings or wide central pits may appear in still larger craters.

These different forms of crater collapse, along with the multiple rings discussed in the next chapter, begin almost immediately after formation of the transient crater. The timescale of collapse is similar to that of excavation, occupying an interval of a few times  $(D/g)^{1/2}$ , the period of a gravity wave of wavelength equal to the crater diameter  $D$  on a planet with surface gravity  $g$ . This interval may range from a few tens of seconds for 10-km diameter craters on the earth to nearly 10 minutes for 300-km diameter craters on the moon. Crater collapse and modification thus take place on timescales very much shorter than most geologic processes.

Collapse of the transient crater results in landforms that have more or less of an interior mass deficiency and more or less of a mass excess around their rims, depending upon the degree of collapse. Ultimate gravitational stability is reached only if the rim continues to subside and the floor rises until a level plain results. This ultimate in gravity-driven modification may be approached by long-term viscous flow of the substrate in which the crater forms, by volcanic flooding and burial of the crater, or by erosional degradation of the rim and infilling of the crater cavity. All of these processes have been observed to occur in different planetary contexts. For example, the nearly level palimpsests on Ganymede and Callisto are attributed to relaxation of craters by viscous flow; many craters on the moon, of

which the Flamsteed ring is the most famous, have been buried by volcanic flows; and the majority of pre-Pleistocene craters on earth have been leveled by erosion. Of these long-term processes, only the less familiar viscous degradation will be discussed in this chapter, on the grounds that volcanic and erosional processes are adequately treated in other texts.

The first portion of this chapter is devoted to a detailed discussion of the collapse that occurs immediately after crater excavation, a process that is peculiar to impact craters. Since the collapse of large craters has never been observed directly, there are many areas of uncertainty in how such features as central peaks, pits, or rings originate. The reader is advised that, in spite of the unified approach presented here, many topics are highly controversial. References to the major alternative points of view are given in the text.

### 8.2 DEBRIS SLIDING IN SIMPLE CRATERS

Simple craters are, by definition, bowl-shaped rimmed pits that superficially resemble the transient craters from which they form. Fresh simple craters on all planetary bodies are morphologically similar, having depth/diameter ratios  $H/D$  between 1/5 (0.2) and 1/3 (0.33). Simple craters range in size from centimeter-scale craters produced in the laboratory or observed by the Apollo astronauts in the lunar regolith, through kilometer-scale craters such as Arizona Meteor Crater on the earth, up to the 90-km-diameter crater Pan on the small Jovian inner satellite Almathea.

Simple craters differ from transient craters in that they are floored by a lens or pool of broken rock (breccia) mixed with shocked debris and impact melt that drains back into the transient crater following excavation. A portion of the rim of the transient crater also slides into the crater cavity. Boulders from the rim of lunar craters are frequently seen to have rolled down into the crater leaving a trail of gouges behind on the crater wall. The inner rims of these craters stand near the angle of repose (about 30°, regardless of surface gravity), consistent with their having been sculptured by massive debris slides. The slope of the crater interior gradually lessens as the center is approached. The center itself is often occupied by a small flat floor.

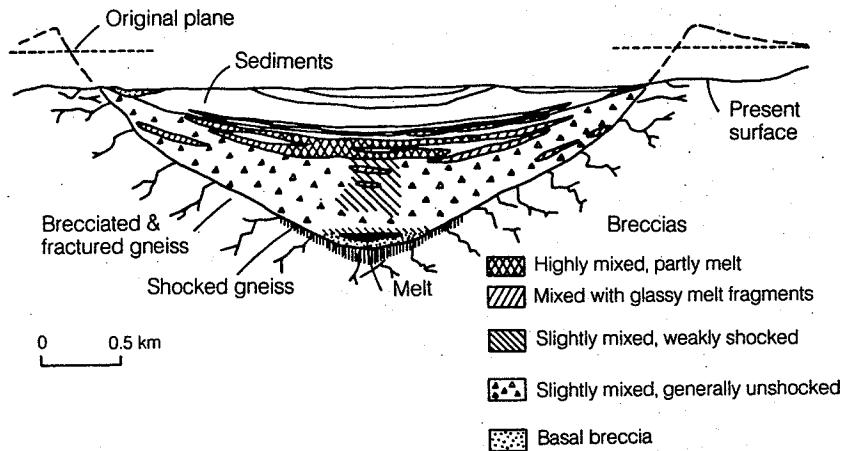
The maximum thickness of the breccia lens is roughly half of the rim-to-floor depth  $H$  of the crater itself. Its volume is about half the volume of the crater, ranging from about 0.4 of the crater volume in the 3.4-km diameter terrestrial Brent crater to 0.7 of the crater volume in 1.0-km diameter Meteor Crater. The maximum depth of brecciation is nearly one-third of the rim diameter  $D$ , based on geophysical investigation of numerous terrestrial craters and one lunar crater, and on direct drilling in a few terrestrial craters. The breccia porosity ranges up to 10 percent, corresponding to a density difference  $\Delta\rho$  between the brecciated rock and the country rock of approximately  $150 \text{ kg/m}^3$ . Bouguer gravity anomalies measured over both lunar and terrestrial craters indicate a total mass deficiency  $\Delta M \approx \pi/24 \Delta\rho D^3$  (Dvorak, 1979), where  $\Delta\rho$  is approximately twice as great on the moon as it is on the earth, which is presumably due to gravitational effects on the brecciation and bulking of the bedrock surrounding the crater.

Laboratory impact craters in sand up to a few tens of centimeters in diameter frequently lack any sign of a breccia lens, although a small quantity of crushed and shocked sand does form in place at the bottom of the crater. On the other hand, craters produced by missile impacts at White Sands Missile Range, New Mexico, which are typically one or more meters in diameter, all contain breccia lenses (Moore, 1976). It is unclear why the laboratory craters lack significant quantities of breccia. One suggestion is that the breccia lens is a result of atmospheric interference with the ejection of debris from the crater, and that laboratory craters lack a breccia lens because they are generally formed in a vacuum. However, in situ investigation of one small crater by the Apollo 17 crew and the gravity anomalies observed over

many other lunar craters from orbit confirm that lunar craters possess breccia lenses comparable to those of terrestrial craters, thus ruling out the atmospheric interference hypothesis.

Figure 8.1 is a cross-section of Brent Crater in Ontario, Canada. Originally about 3.4 km in diameter, it contains an estimated  $2.1 \text{ km}^3$  of breccia fill. Of this, about  $0.055 \text{ km}^3$  was melt, in good agreement with the melt scaling relation (Eq. 7.10.1a). The melt rocks are not uniformly distributed throughout the breccia, however. The cross section shows a small, nearly homogeneous pool of solidified melt that contains few clasts near the base of the breccia lens. The main mass of breccia overlying this melt pool is only weakly shocked and nearly homogeneous in composition. More solidified melt occurs near the top of the breccia lens where it is mixed with an abundance of clasts from the bedrock. These clasts show a wide range of shock levels. Similar structures have been reported from Meteor Crater, Lonar Lake Crater (1.8 km in diameter) in India, and West Hawk Lake crater (3 km in diameter) in Manitoba, Canada. Solidified melt pools are not found at Meteor Crater or Lonar Lake, but a basal zone of highly shocked breccia is invariably present, indicating similar conditions at the bottom of the breccia lens. This bottom zone of highly shocked material is generally accepted to be the bottom of the transient crater. The absence of melt in these craters may have been the result of a subsurface water table which dispersed the melt by a steam explosion.

The meaning of this peculiar distribution of highly shocked or melted debris was clarified in 1977 by Canadian geologists R. A. F. Grieve, M. R. Dence, and P. B. Robertson. The most highly shocked rocks originally form near the crater's center in immediate proximity to the impactor.



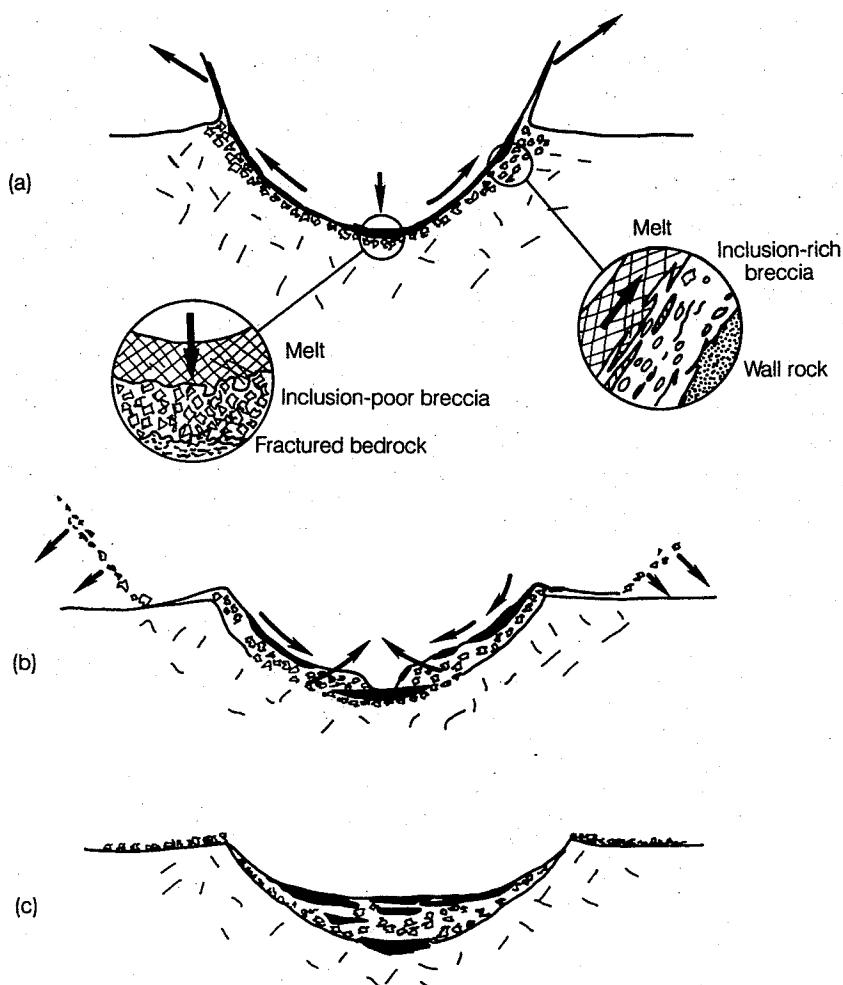
**Fig. 8.1** Geologic cross section of Brent Crater in Ontario, Canada. Brent is a well-studied simple crater that formed in crystalline rocks. Although the rim is eroded away, drilling revealed that the crater is filled with lightly shocked breccias. A small melt pool occurs at the base of the breccia lens and more heavily shocked rocks occur near its top. After Grieve et al. (1977).

As the crater expands during excavation the shocked and melted rocks line the interior of the growing crater cavity. Near the center of the crater these rocks are driven nearly straight downward and mix only slightly with the underlying, less-shocked breccia. The melt near the crater walls, however, has a large velocity component tangential to the walls (see Sec. 5.5 on the excavation flow field). The melt and breccia in this zone is subjected to strong shear, and considerable mixing may take place, contaminating the melt flowing along the walls with clasts from the less-shocked breccia. In spite of this mixing, however, at the end of the excavation stage the debris in the transient crater still retains its original shock stratigraphy with melt overlying breccia, which in turn overlies fractured bedrock. This structure is illustrated in Figure 8.2a.

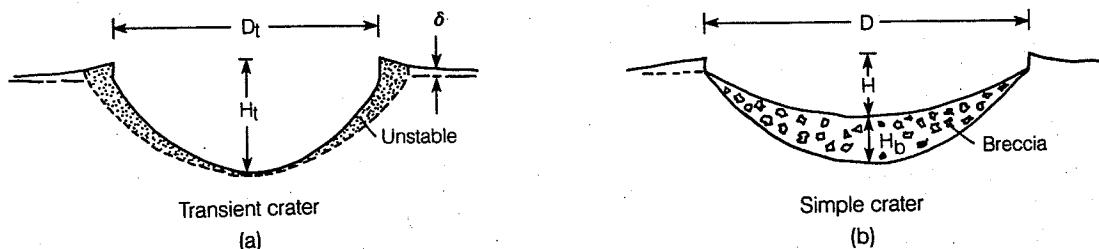
The debris lining the interior of the transient

crater ceases its upward flow at the end of the excavation stage and begins to collapse inward as modification begins. The oversteepened rim of the transient crater slides downward, leaving behind a slope standing nearly at the angle of repose. Meanwhile, a slip surface (or surfaces) develops near the toe of the slumping debris, allowing weakly shocked breccia to override and bury the central zone of nearly homogeneous melt, as in Figure 8.2b. Debris slumping centripetally from the walls meets in the center and rapidly covers the transient crater floor. The last material to slide onto the floor is the clast-rich melt rock that originally lined the transient crater rim. This material forms the extensive sheets of mixed breccia that are observed near the top of the breccia lens, shown schematically in Figure 8.2c.

The extent of the collapse and the dimensions of the original transient crater may be estimated



**Fig. 8.2** Formation of the breccia lens in a simple crater. The growing transient crater (a) is lined with breccia and melt that flows upward and out of the crater along its walls. Collapse begins once the rim has formed (b) and the breccia slides back into the crater. The breccia sliding from the walls overrides the melt on the transient crater's floor and traps a melt pocket. At the cessation of sliding (c), the breccia lens has buried a deep melt pocket. The remains of the mixed melt and breccia that lined the transient crater are now concentrated near the top of the breccia lens.



**Fig. 8.3** A geometric model of simple crater formation. Based on volume conservation, the model relates the depth  $H_t$  and diameter  $D_t$  of the transient crater to the depth  $H$ , diameter  $D$ , and breccia lens thickness  $H_b$  of the final simple crater. The rim thickness  $\delta$  is taken into account by this model.

from the simple model illustrated in Figure 8.3. In this model the shape of both the transient and final simple crater is represented by a paraboloid of revolution: that is, the difference in altitude  $z$  between a point on the crater wall and the lowest point in the crater center is given by  $z = 4H(r/D)^2$ . The height of the rim above the preimpact surface is given by  $\delta = H_t/40(D_t/r)^3$ , where the subscript  $t$  denotes quantities describing the transient crater. These formulae are in close agreement with the shapes of observed craters. The constants are chosen so that the rim above the preimpact surface is equal in volume to the crater cavity below the preimpact surface. No provision is made for bulking of ejecta.

When the transient crater collapses, material from both the rim and the crater wall slides onto the floor of the crater cavity. Equating the volume of the breccia lens to the volume of this debris yields an equation for the transient crater diameter:

$$D_t = \left( 1 - \frac{5}{4} \frac{H_b}{(H + H_b)} \right)^{1/3} D \quad (8.2.1)$$

where  $H_b$  is the maximum thickness of the breccia lens and the depth of the transient crater  $H_t$  is assumed to equal  $H + H_b$ . The right-hand side of Equation 8.2.1 is written only in terms of observable quantities. Substitution of the observed breccia lens thickness,  $H_b \approx H/2$ , gives an equation for the transient crater diameter  $D_t = 0.84D$ . The final simple crater is thus about 19 percent larger than the initial transient crater. The depth/diameter ratio of the transient crater is similarly estimated to be  $H_t/D_t \approx 1/2.7$ . These figures show that the final simple crater is only slightly larger in diameter than the transient crater but is significantly shallower. The above model also predicts that the rim height  $h_R$  of a simple crater should be  $0.035D$  for a final depth/diameter ratio of 1/5, in remarkable agreement with observed rim heights, Equation 6.2.1.

The timescale of the transient crater's collapse can be estimated from elementary mechanics. Suppose the debris slides without friction down an inclined plane with average slope  $\theta$ , equal to

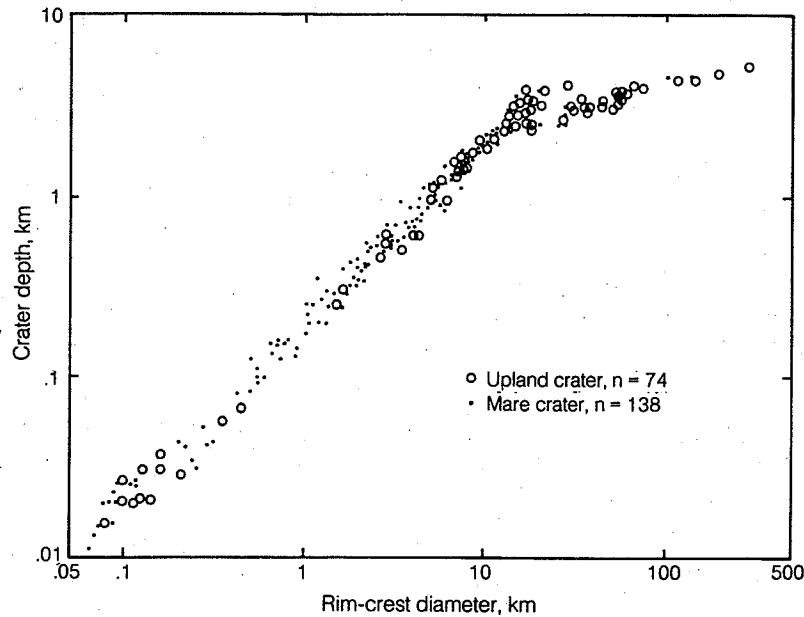
the average internal slope of the transient crater of depth  $H_t$  and diameter  $D$ . Thus,  $\tan \theta = 2H_t/D$ . The time required for a portion of the debris to traverse this slope is  $(2H_t/g \sin \theta)^{1/2}$ . For small angles  $\theta$ ,  $\sin \theta \approx \tan \theta$ . The collapse timescale is thus of order  $(D/g)^{1/2}$ , as asserted in Section 8.1. Of course the debris does not move without friction, so the actual timescale should be somewhat longer. On the other hand, the debris does not have to traverse the entire length of the slope, so the actual timescale should be somewhat shorter. These effects thus tend to cancel, and the elementary estimate is probably reasonably accurate.

The velocity of the debris as it reaches the center of the crater can be similarly estimated to be on the order of  $(2gH_t)^{1/2}$ . This velocity may reach 100 m/second in 15-km diameter lunar craters, comparable to velocities observed in large terrestrial rock avalanches. It should come as no surprise that these fast-moving debris streams often pile up in low mounds when they meet in the center of the crater, raising central peaks and hummocks on the crater floor. Such central peaks in the debris lens, however, are quite different from the central peaks observed in large complex craters such as Copernicus. The latter type of central peak is due to wholesale uplift of the crater floor, described in the next section.

The melt buried in the breccia lens of even simple craters may have important consequences for the thermal structure of the impacted surface. The total mass of melted target rock may easily exceed several projectile masses, depending upon the impact velocity and target material (see Eq. 7.10.1a and Fig. 7.5). This melt, along with the associated shocked and heated rocks, contains the bulk of the energy deposited by the projectile (see Sec. 5.2). Much of this hot rock is buried at depths comparable to the final crater depth in simple craters. The near-surface melt layers cool relatively quickly, but the cooling time of the buried melt pockets is on the order of  $H^2/\kappa$ , where  $\kappa$ , the thermal diffusivity, is typically  $1 \times 10^{-6} \text{ m}^2/\text{second}$  for rock. The cooling time for melt in the breccia lens underlying a 15-km diameter crater is thus about 100,000 years.

This thermal perturbation contributes signifi-

## IMPACT CRATERING



**Fig. 8.4** Depth/diameter relation for fresh lunar craters. Craters smaller than about 15 km in diameter have depths equal to about one-fifth of their diameter. The depths of larger craters are considerably smaller, indicating a major modification of their structure. Craters with depths proportional to diameter are simple, whereas craters on the shallow depth trend are complex, with central peaks and slumped walls. There is a small overlap of the two populations near the transition diameter. *After Pike (1974).*

cantly to the heat flow near the crater for a long period after the impact. During planetary accretion and the period of heavy bombardment early in solar system history this buried heat may have contributed importantly to planetary thermal budgets. Slowly cooling melt rocks may also produce both igneous-textured, coarse-grained "plutonic" rocks and thermal metamorphism of adjacent rocks, both on major planets and on asteroids that lack internal heat sources. It has also been suggested that hot rock in the breccia lenses of Martian impact craters might induce local hydrothermal circulations, forming clays and other weathering products and perhaps initiating catastrophic floods.

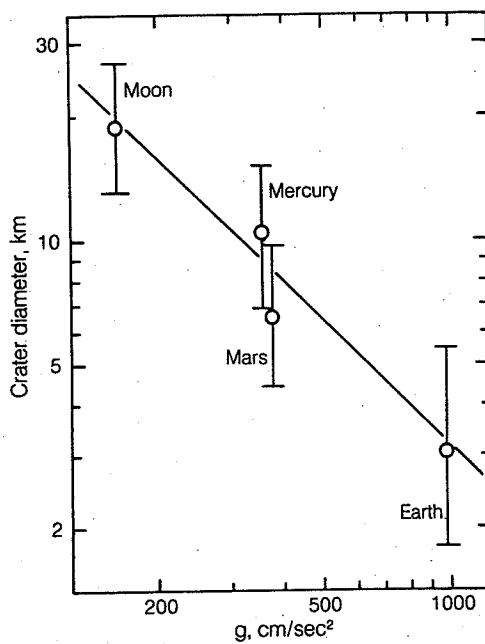
### 8.3 COMPLEX CRATERS

The larger an impact crater is, the further the initial transient crater departs from gravitational stability. Simple craters are the result of a relatively small-scale collapse of the steep crater rim. Craters larger than a well-defined threshold size collapse more spectacularly, completely altering the appearance of the transient crater, producing final craters with central peaks, terraced walls, and, at the larger sizes, flat floors. Such craters were named "complex craters" by M. R. Dence in 1965. The process of collapse is known as *slumping*.

One of the most remarkable features about the simple-to-complex transition is its abruptness.

Figure 8.4 illustrates this on a plot of crater depth as a function of diameter for a sample of 212 fresh, unflooded lunar craters. Crater depth is directly proportional to rim-crest diameter for all craters smaller than about 15 km in diameter. These craters are all of the simple type with a depth/diameter ratio close to 1/5. This relation changes suddenly for craters larger than about 15 km, marked by a pronounced "knee" in the plot. Large crater depths increase only slowly with increasing diameter, ranging from 3 km deep for craters about 20 km in diameter to 5 km deep for craters exceeding 300 km in diameter. These large craters all have the "complex" morphology, including central peaks, flat floors, and terraced walls. A small population of craters lies along the extension of the simple crater line in the 15- to 20-km diameter transition region. These craters have the simple morphology, whereas craters falling along the large crater line in this transition region have the complex morphology. This slight overlap supports the idea that transient craters evolve into either the simple or complex form, depending upon size, except in the narrow transition region where local factors may influence which morphology develops.

The simple-to-complex transition is almost undoubtedly a consequence of collapse under gravity when some strength threshold is exceeded. The exact nature of this threshold will be discussed subsequently. The importance of gravity is clear from Figure 8.5, in which the crater diameter at



**Fig. 8.5** Crater diameter at the simple-to-complex transition on the earth, moon, Mercury, and Mars. The transition is a function of  $1/g$ , where  $g$  is the surface acceleration of gravity. The transition on Mars is at a somewhat smaller diameter than that for the other planets, suggesting that volatiles, which account for the fluidized ejecta blankets around Martian craters, may also enhance the transition to the complex crater form. *Data from Pike (1988).*

which the simple-to-complex transition occurs is plotted against surface gravity  $g$  for Mars, Mercury, the earth, and the moon. In spite of large uncertainties, the transition diameter clearly follows a  $1/g$  dependence. The transition diameter on Mars falls slightly below the curve, but since Martian crater ejecta deposits are often fluidized by the inclusion of some volatile substance, probably liquid water (see Sec. 6.3.5), it is likely that crater collapse is enhanced by this substance. The Mercurian point in Figure 8.5 has been somewhat controversial. Many investigators found that it lies significantly above the  $1/g$  line. Very recent work, however, now shows that it really does lie on the line (Pike, 1988). Preliminary data on Venusian craters from Veneras 8 and 9 is consistent with a  $1/g$  dependence, but is not yet accurate enough to plot as a separate point. The transition diameter on the icy satellites of Jupiter and Saturn lies significantly below the line defined by the terrestrial planets, suggesting that ice is in some sense weaker than the rocky surfaces of the terrestrial planets.

The depth of the deepest simple crater at the onset of the simple-complex transition is also a function of  $1/g$ . Thus, the deepest simple crater on the moon is about 3 km deep, whereas on Mer-

cury, Mars, and the earth the deepest simple crater is 1.7 km, 1 km, and 0.5 km deep, respectively.

The following sections examine the morphology of complex craters and discuss the various interpretations and theories of their features.

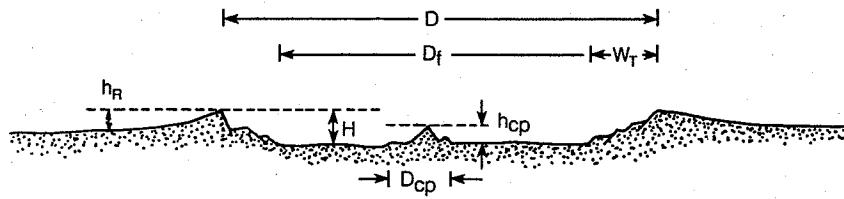
### 8.3.1 Morphology and morphometry of complex craters

Complex craters constitute one of the most common large-scale surface features found on airless bodies throughout the solar system. They have been studied intensively since before the era of spacecraft exploration of the moon and planets. Several examples of complex craters are even known on the earth in various states of exposure and preservation. Nevertheless, many aspects of their structure and formation are either unclear or controversial. The nature and origin of the prominent central peaks is still argued heatedly. The collapse that seems obviously necessary to account for complex crater formation violates the standard notions of material strength of geologic materials. The origin of multiple rings, both internal and external, has nearly as many interpretations as there are investigators.

Any investigation, however, begins with a description of the object to be investigated. Much effort has been expended on the quantitative description of complex craters, beginning with R. Baldwin's classic 1949 book, *The Face of the Moon*. Modern practitioners of this art now call the field *morphometry*, as opposed to the more qualitatively descriptive *morphology*, and I conform to this practice here. Modern work also appeals to large samples, careful statistical analysis, and, most important, interplanet comparisons. Some results of this type of work on depth/diameter ratios have already been referred to in the introduction to this section, but complex craters offer many other measurable dimensions.

Figure 8.6 is a schematic cross section of a "typical" complex crater, modeled mainly after such lunar craters as Copernicus (Fig. 6.10) or Theophilus (Fig. 2.6). The figure illustrates the principal features of a complex crater: a single or group of central mountains, surrounded by a relatively flat floor, inside of a terraced rim. The rim itself and the ejecta beyond it are treated extensively in Chapter 6. The summits of the central peaks occasionally approach the elevation of the flat plain outside the crater, but, on the moon at least (there may be exceptions on Mars), they are seldom higher than the crater rim. The flat floor, so conspicuous in 100-km-diameter lunar craters, is often absent in the smallest complex craters. The number of terraces and the extent of the terraced zone ranges from a single terrace in the smallest craters to dozens in larger craters, in which the zone may be tens of kilometers wide. Figure 8.6

## IMPACT CRATERING



**Fig. 8.6** Principal morphometric parameters used to describe complex craters.  $D$  is the rim-crest diameter,  $D_f$  the diameter of the flat inner floor,  $W_T$  is the width of the terrace zone ( $w$  is the width of the largest terrace),  $H$  is the crater depth,  $h_R$  its rim height,  $D_{cp}$  is the diameter of the central peak complex and  $h_{cp}$  is the height of the central peaks above the crater floor. For the "typical" lunar crater Copernicus,  $D = 93$  km,  $D_f = 54$  km,  $W_T = 19.5$  km,  $H = 3.8$  km,  $h_R = 1.1$  km,  $D_{cp} = 25$  km, and  $h_{cp} = 2$  km.

also defines a number of commonly used dimensional measures of complex crater geometry.

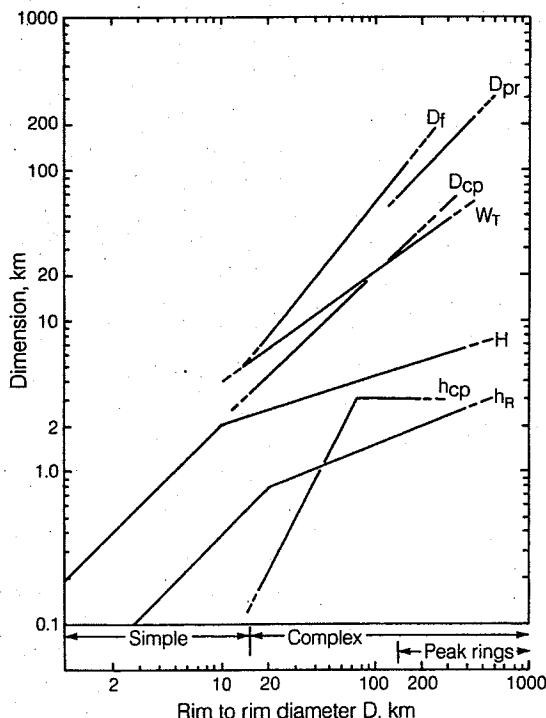
The results of many years of investigation of fresh lunar craters are summarized in graphical form by Figure 8.7 and in parametric equations by Table 8.1, where each pertinent dimension is given as a function of the final crater's rim-to-rim diameter. The straight lines for crater depth,  $H$ , in Figure 8.7 reproduce the lunar depth/diameter relation, Figure 8.4. Most dimensions increase monotonically with rim diameter, although the rate of increase varies. Central peak height,  $h_{cp}$ , however, increases only up to a point. The height of the central peaks above the crater floor at first

increases rapidly with crater diameter beyond the simple-to-complex transition. This increase stops, however, once the crater exceeds about 80 km in diameter, and the peak height remains nearly constant through 140-km diameter craters, beyond which central peaks disappear.

A ring of mountains first develops within the crater at about 140 km in diameter. Some craters, such as Compton (Fig. 8.8a) and Antoniadi on the moon, possess both a central peak and a broad ring of mountain summits. The diameter of this inner "peak ring" is roughly half the crater rim diameter, not only for lunar craters but for Mercurian and Martian craters as well (Fig. 8.9). The central peak diameter,  $D_{cp}$ , is itself a constant fraction  $0.22 \pm 0.03$  of the crater diameter (Pike, 1985), which is also apparently independent of the planet on which the crater forms. This includes the moon, Mercury, Mars, and the earth (see Fig. 8.10). It also may include craters on Ganymede and Callisto. The old idea that the central peaks shrink as the transition to peak ring craters approaches evidently applies to central peak volume, but not diameter.

Complex craters larger than about 140 km in diameter on the moon lack central peaks. Instead, their floors are occupied by peak rings. The central area of these craters is a relatively flat plain encircled first by the peak ring and then by the crater rim. In these craters the rim is recognizable as such by the terraces that dissect its interior wall. The lunar crater Schrödinger (Fig. 8.8b) is typical of this class, but peak ring craters are also well developed on Mars (Fig. 8.8c) and Mercury (Fig. 8.8d). Although the diameter of the peak ring seems to be a constant fraction (0.5) of the crater diameter, independent of planet, the onset of peak ring development does depend on the planet's gravitational field (see again Figure 8.9). Thus, the transition from central peaks to peak rings occurs at a diameter of 140 km on the moon, 75 km on Mercury, 45 km on Mars, and at less than 25 km on the earth (Pike, 1983), in the same kind of  $1/g$  proportion—with the same exception for Mars—as observed for the simple-to-complex transition.

Another variant of the complex crater type is



**Fig. 8.7** Morphometry of fresh lunar craters. The dependence of the parameters defined in Figure 8.6 and Table 8.1 on crater diameter is illustrated. Also shown is peak-ring diameter,  $D_{pr}$ , for craters larger than about 140 km in diameter.

TABLE 8.1 *Lunar crater morphometry*

Parameter	Dependence on Rim-to-Rim Diameter ( $D$ , km)	Diameter Range (km)	Source
Crater Depth	$H = 0.196 D^{1.01}$ $H = 1.044 D^{0.301}$	<11 11-400	*
Crater Floor Dia.	$D_f = 0.19 D^{1.25}$	20-140	*
Central Peak Dia.	$D_{cp} = 0.22 D$	20-140	†
Peak Ring Dia.	$D_{pr} = 0.50 D$	140-450	‡
Central Peak Height	$h_{cp} = 0.0006 D^{1.97}$ $h_{cp} \approx 3$	15-80 80-200	**
Rim Height	$h_R = 0.036 D^{1.014}$ $h_R = 0.236 D^{0.399}$	<21 21-400	*
Terrace Zone Width	$W_T = 0.92 D^{0.67}$	15-350	††
Widest Terrace Width	$w = 0.09 D^{0.87}$	20-200	‡‡

\*Pike (1977)

†Pike (1985)

‡Wood and Head (1976)

\*\*Hale and Grieve (1982)

††Based on data from Pike (1976)

‡‡Pearce and Melosh (1986)

well developed on Ganymede and Callisto, although it is also occasionally observed on Mars and a few lunar examples are claimed. This is the central *pit* crater (Fig. 8.11). Central pits are present in craters on Ganymede that range from 16 to more than 120 km in diameter. Although rare at the smaller crater sizes (most craters larger than 5 km in diameter have central peaks), essentially all craters larger than 30 km in diameter on both Ganymede and Callisto have pits. Pitted craters are generally seen to have terraced walls, when resolution is adequate to resolve terraces, and hence show evidence for collapse. The pits are not mere depressions in the crater floor, but are frequently rimmed and in larger craters show up-bowed floors, although these features are probably due to viscous relaxation (see Sec. 8.4).

The pit diameter generally increases with crater diameter, although there is a great deal of scatter in the relation. One reason for this scatter is an apparent age dependence in which older craters have pit diameters larger than those in younger craters. This can presumably be attributed to a difference in crustal properties with the geologic age of the cratering event. Pit diameter is not, however, a linear function of crater diameter. The best fits to the data relating pit diameter  $D_p$  to crater diameter  $D$  for Ganymede (Passey and Shoemaker, 1982) give:

$$D_p = 1.9 \exp(0.023D) \quad (8.3.1)$$

and for Callisto:

$$D_p = 1.45 \exp(0.028D) \quad (8.3.2)$$

where  $D_p$  and  $D$  are both in kilometers.

Pit craters on Ganymede and Callisto appear to occupy the same relative position in the diameter-morphology sequence as peak rings do on the

terrestrial planets. Only a few possible peak ring basins have been identified on either icy satellite, suggesting that pit formation somehow precludes or supersedes peak ring formation. The central pits, however, are much smaller in diameter than peak rings would be in the same size crater, so it seems unlikely that any simple connection between the two crater forms exists.

A number of explanations have been offered for the formation of central pits. Although many center on the unique properties of ice and its relative ease of melting or vaporization, one suggests that central pit craters are the result of impacts of unusually low velocity and that the pit is the trace of the projectile that produced the crater. None, however, has yet reached the point of general acceptance and the origin of pit craters remains, for the present, mysterious.

### 8.3.2 Geologic structure of complex craters and central peaks

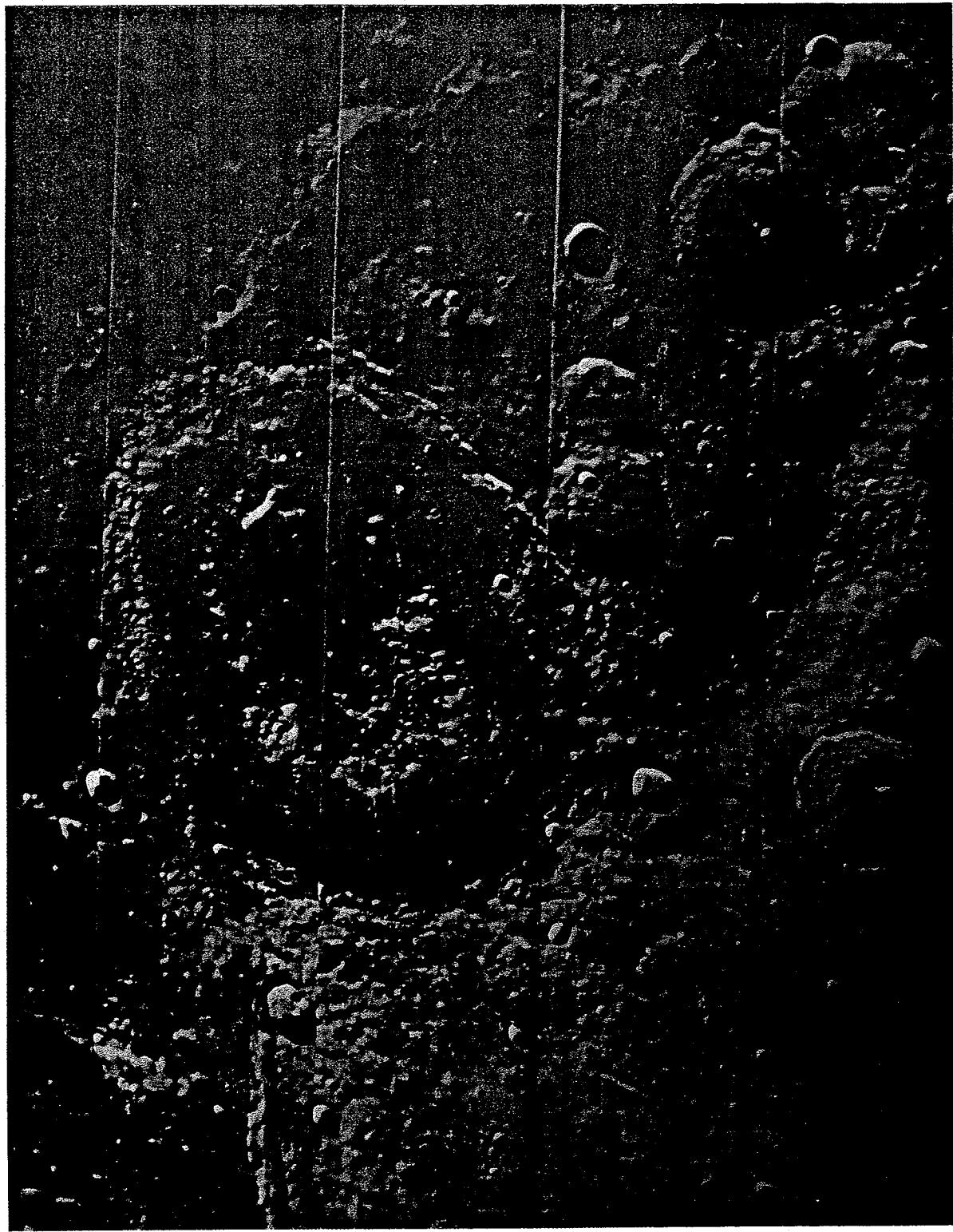
Although the morphometric data discussed in the preceding section provides useful constraints on the processes by which complex craters form, geologic data provide much stronger constraints on the subsurface motions that result in the observed final structure. Unfortunately, geologic cross sections are not available for any extraterrestrial crater. Only the lunar crater Copernicus afforded a modicum of geologic information after C. Pieter's (1982) spectrophotometric discovery of olivine-bearing rock outcrops among its central peaks. These rocks are presumed to have been uplifted from deeper levels within the moon's crust. A great deal of our understanding of complex crater formation thus rests upon observations and interpretations of terrestrial craters. In using terrestrial



Fig. 8.8 Peak-ring craters on the moon, Mars, and Mercury. (a) The 175-km diameter crater Compton on the moon is transitional between central peak and peak-rings craters; it possess one of each. *Lunar Orbiter Photograph LO V 181M.* (b) Schrödinger is a 320-km diameter peak-ring crater on the moon. *Lunar Orbiter Photograph LO IV 94M.* (c) The 220-km diameter crater Galle is a Martian peak-

*Lunar Orbiter Photograph LO V 181M.* (b) Schrödinger is a 320-km diameter peak-ring crater on the moon. *Lunar Orbiter Photopgraph LO IV 94M.* (c) The 220-km diameter crater Galle is a Martian peak-

(b)



ring crater, seen obliquely. *Viking mosaic 211-5428.* (d) Mercurian craters Strindberg ( $D = 165$  km) to the lower left and Ahmad Baba ( $D = 115$  km) to the upper right are both peak-ring craters. *Mariner 10 FDS 150, rectified.*

(c)



data to interpret the formation of extraterrestrial craters, we must hope that the special circumstances of terrestrial environment, particularly earth's atmosphere, are not of overriding importance in complex crater formation.

Study of simple craters might suggest that complex crater collapse is merely an exaggerated development of the debris sliding that creates the breccia lens in simple craters. Indeed, large simple craters often possess low central or near-central mounds that some workers have identified as protocentral peaks on the basis of morphology. These mounds form within the mass of breccia filling the lower part of the crater and, as discussed in Section 8.2, are probably the result of the convergence and pileup of high-speed debris streams sliding down the walls and onto the crater floor. Meteor Crater, Arizona, for example, possesses a 15-m-high central mound of this type that is displaced to the northwest of the crater center.

Geologic investigation of the central uplifts of terrestrial complex craters, however, shows that they are composed of deformed and fractured rocks that originally underlay the transient crater. They have been stratigraphically uplifted distances comparable to the transient crater depth. Although the rock composing these uplifts is fractured, its original stratigraphy (if it had one) can usually be discerned: it is *not* a mixed breccia of the type occurring below the floors of simple craters. The geologic structure of these craters thus reveals a more profound difference between simple and complex crater types than that suggested by morphology alone.

Figure 8.12 is a geologic cross-section of the 13-km diameter Sierra Madera structure in western

Texas. The age of this deeply eroded complex structure is either late Cretaceous or early Tertiary. The central uplift is about 5 km in diameter. Rock units in this uplift have been raised approximately 1.2 km from their preimpact positions. Thickening and folding of rock units in the uplift implies additional inward displacements of several hundred meters. A 1- to 3-km-wide ring depression surrounds the uplift. The depression is enclosed by a ring of concentric normal faults, downthrown toward the center, that presumably represent the roots of the terraced zone and are thus close to the position of the crater's preerosion rim. Erosion has completely erased the original topographic form, along with any vestige of ejecta or melt within the crater.

Careful study of Sierra Madera and 14 other terrestrial complex craters ranging between 4 and 100 km in diameter in both sedimentary and crystalline rocks (Grieve et al., 1981) has shown that the stratigraphic uplift of the crater's center,  $h_{su}$ , is related to the final crater diameter  $D$  by

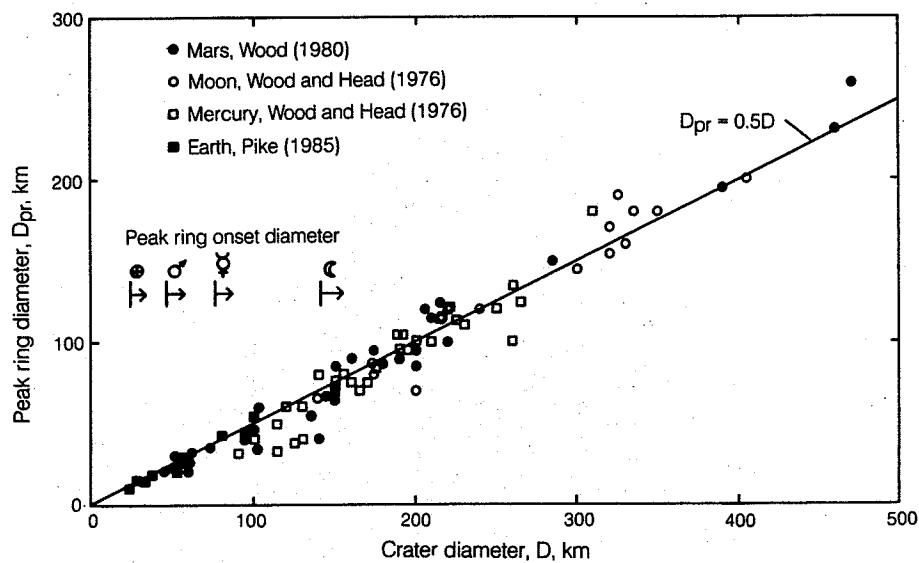
$$h_{su} = 0.06D^{1.1} \quad (8.3.3)$$

where all distances are in kilometers. The depth of the transient crater,  $H_t$ , is roughly twice as large as the stratigraphic uplift when account is taken of the initial downward deflection of strata beneath the transient crater (see Sec. 5.5). Horizontal strata are thus first pushed downward by the excavation flow, then uplifted by  $h_{su}$  above their initial depths. Taken at face value, this seems to imply that the transient crater has a depth/diameter ratio of only about 0.16. However, the transient crater diameter  $D_t$  is itself only a fraction of the final crater diameter  $D$ . Examination of the

(P)



## IMPACT CRATERING

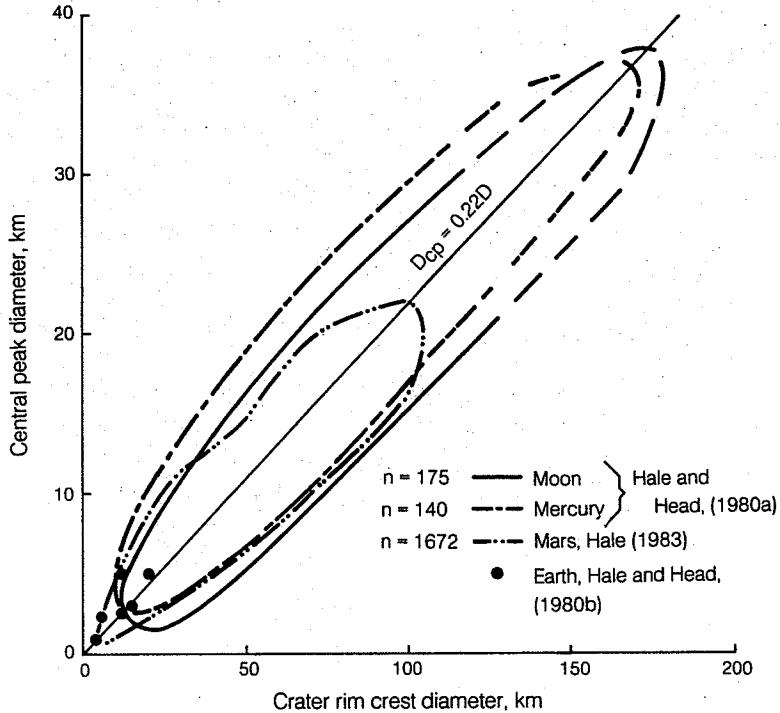


**Fig. 8.9** The relation between peak-ring diameter and crater rim-crest diameter is universal on the terrestrial planets. Only the onset diameters vary with the same inverse dependence on surface gravity seen for the simple-to-complex transition. The onset diameter on Mars is somewhat smaller than for Mercury, also similar to the simple-to-complex transition.

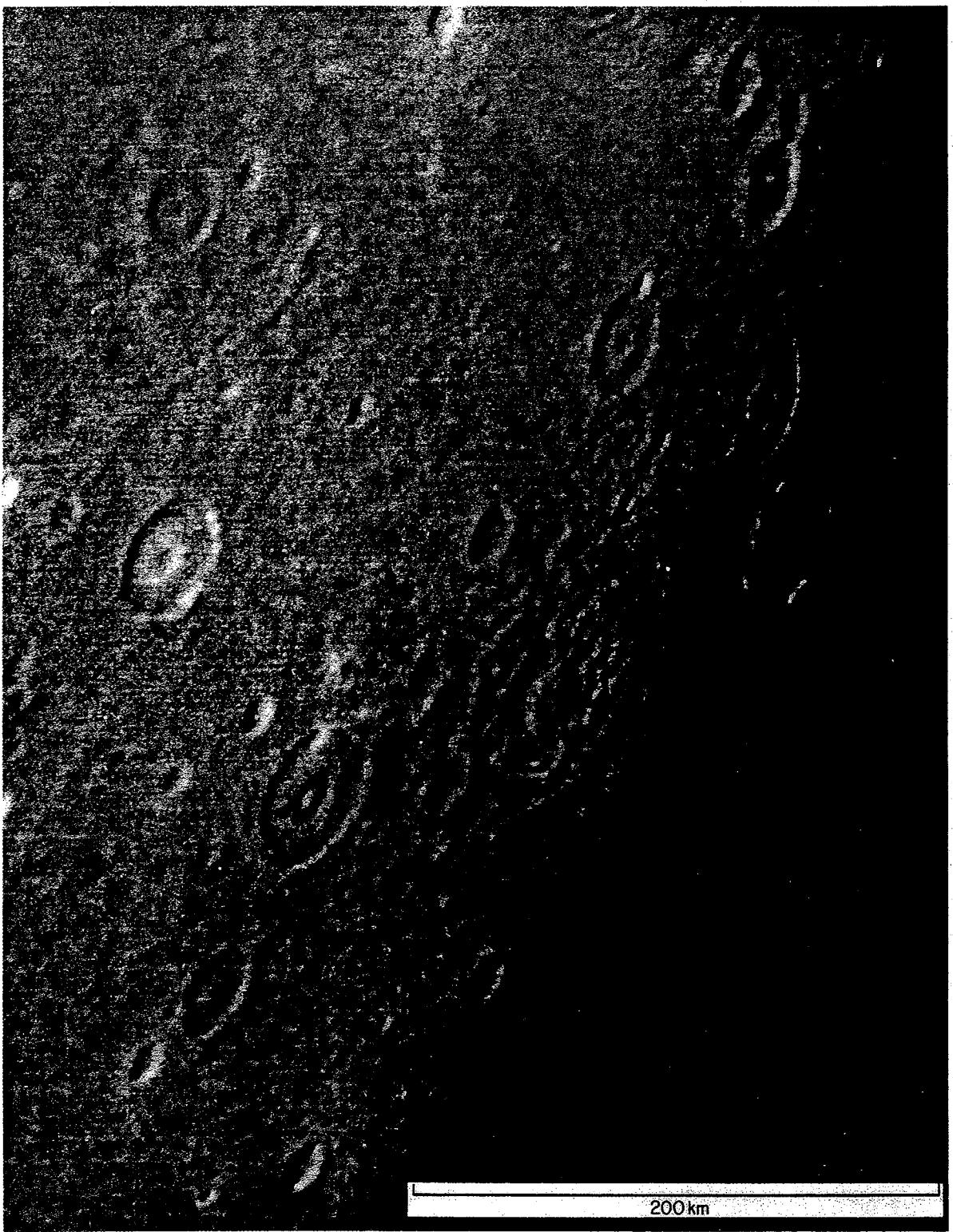
limit of deep excavation in these craters suggests that  $D_r \approx 0.5$  to  $0.65D$ , so the transient crater depth/diameter ratio is roughly 0.24 to 0.32, in excellent agreement with simple craters and small-scale impact experiments.

These large stratigraphic uplifts are also supported by the orientation of shatter cones in the

rocks surrounding the crater (see Sec. 3.4.2 for a description of shatter cones and Table 3.2 for the pressures at which they form). Shatter cones originally form with their axes pointing toward the source of the shock wave that created them. Detailed mapping of shatter cone orientations in complex craters, however, shows that they point

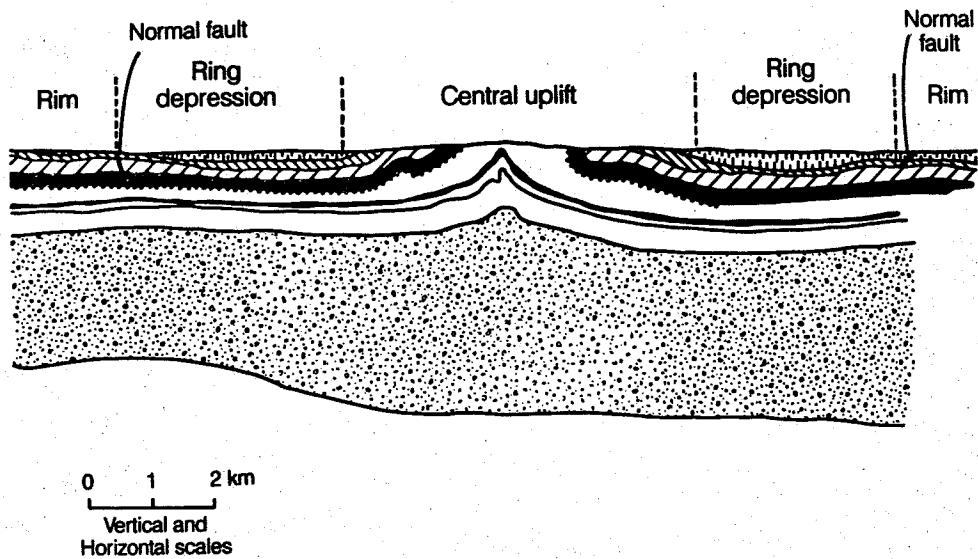


**Fig. 8.10** The central peak diameter in complex craters is roughly 0.22 of the rim-crest diameter on all of the terrestrial planets, although the individual scatter is large.



**Fig. 8.11** Central pit craters on Ganymede. The craters near the terminator show anomalously large central pits and upbowed floors. The upbowed floors are probably a result of viscous relaxation subsequent to crater excavation. *Voyager FDS 20631.33*.

## IMPACT CRATERING



**Fig. 8.12** Simplified geologic cross section of the Sierra Madera impact structure in west Texas. Sierra Madera appears to be a terrestrial complex crater. Although the surface rocks have been eroded away, the cross section shows a well-developed central uplift surrounded by a ring-shaped depression bounded by normal faults. These faults may correspond to the terraces seen in fresh lunar craters. *Simplified from Wilshire and Howard (1968). Copyright 1968 by the AAAS.*

too steeply upward (or even point outward) in the central region of the crater. The simplest explanation of this observation is that the rock units were uplifted and tilted away from the crater center following the passage of the shock wave. Reconstructions of the original crater shape from shatter cone orientations agree with the stratigraphic interpretations in requiring that the crater originally had the form of a deep bowl with a depth/diameter ratio similar to that inferred for transient craters.

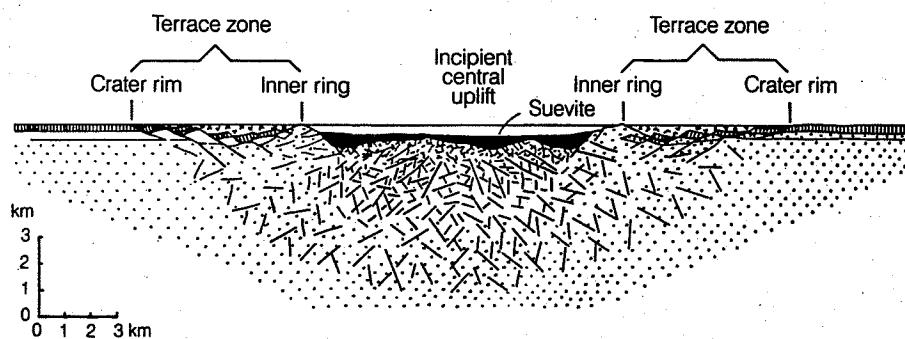
Whether the depth/diameter ratio of transient craters is actually independent of crater size is currently uncertain. Although there is some evidence for a decrease of  $H/D$  with increasing velocity or size, the evidence is equivocal. Detailed geologic reconstructions of preslump lunar crater profiles (Settle and Head, 1979) suggest that any decrease of  $H/D$  with crater size occurs only in craters larger than 70 to 100 km diameters. The uncollapsed 90-km-diameter crater Pan on the small Jovian satellite Almethea is between 10 and 20 km deep, also supporting a constant depth/diameter ratio for large craters.

A peculiarity of many terrestrial complex craters is a shallow excavated shelf that often surrounds the central region of deep excavation. The excavated volume, before the central uplift fills it, has the form of an inverted sombrero with a deep parabolic crown and shallow broad brim. Whether this form of excavation is also characteristic of extraterrestrial craters is problematic. However, the shelf could have been excavated by an airblast accompanying the impact or by frag-

ments aerodynamically stripped from the impactor as it plunged through the earth's atmosphere (see Sec. 11.2.4). There is thus some reason to suspect that this aspect, at least, of terrestrial complex crater structure may not apply to craters formed on airless bodies.

Ries Crater in Germany appears to be a terrestrial peak-ring crater. A simplified cross-section of this 22-km diameter structure is presented in Figure 8.13. The 9-km-diameter central depression is filled with up to a 400-m thickness of Suevite (mixed glass and shocked lithic fragments). Surrounding this depression is a ring of uplifted and fractured basement, beyond which are blocks bounded by normal faults that carry sedimentary cover rocks downward into the crater. Most of these blocks are overlain by ejecta. The crystalline rocks beneath the Ries crater are shattered and broken to a depth of at least 6 km, inferred from the low seismic velocity beneath the crater and an 18-mgal gravity anomaly over it.

Structural studies of terrestrial craters thus leave little doubt that the modification from the bowl-shaped transient crater to the complex form is achieved by wholesale collapse of an initially deep transient crater. This collapse is achieved mainly by uplift of the rocks underlying the crater's center, while rock units near the rim slump downward and inward. Interplanet comparisons of the diameter at the onset of collapse prove that gravity is the principal force driving it. Before going on to examine the mechanics of this collapse in the next section, however, a brief digression on alternative ideas about the origin of cen-



**Fig. 8.13** Geologic cross section of Ries Crater, Germany. Based on both drilling and geophysical data, this cross section suggests that the 22-km diameter Ries Crater is a terrestrial peak-ring basin. Its central basin is filled with Suevite, a mixture of highly shocked and melted rock and cold clasts. Simplified from Pohl *et al.* (1977).

tral peaks is in order, especially since some central peaks do originate by mechanisms other than collapse.

Central peaks may form in impact craters by at least three mechanisms besides collapse. Section 5.6.2 discussed the rather special morphology of craters that form in layered targets where a weak layer overlies a strong layer. Within a limited diameter range a low central mound remains on the crater floor, which coincides with the top of the strong layer. This central mound is an erosional remnant, left behind in the stagnation zone where the downward and outward then upward excavation flow encounters the strong layer and fails to excavate it. Although central peaks of this type occur abundantly within lunar mare craters less than 200 m in diameter, this mechanism clearly cannot explain the central peaks observed in large impact craters since it is independent of gravity and because it predicts no stratigraphic uplift at all.

R. B. Baldwin (1963) suggested that central peaks originate by a "rebound" process in which rocks below the crater are strongly compressed by the force of the impact and then spring back elastically when the stress is relieved, buckling the crater floor upward into a structural dome. Although this process has no obvious dependence on gravity or crater size, and so probably cannot explain the central peaks of complex craters, it may be important in the formation of central peaks in the shallow-buried explosion craters or impact craters where the projectile has an unusually small depth of penetration. Central peaks are commonly observed in surface explosion craters, although they do not occur for buried charges. Detailed numerical studies of a 20-ton TNT surface explosion (Ullrich *et al.*, 1977) suggest that some combination of elastic rebound and gravitational collapse leads to the central mound observed in that experimental test.

The effective depth of burial of impact craters is evidently larger than that required to produce a

rebound central peak. However, in special cases where the depth of penetration is unusually low, rebound central peaks may arise. This may have occurred in the 40-m diameter crater produced by the Apollo 14 Saturn IVB upper stage on the moon. This crater has a central peak, unlike nearby craters of similar diameter, perhaps because of the low depth of penetration of the empty booster, although the peculiar structure of the projectile (a large hollow cylinder of low *average* density) may have had an important effect. Low-density comets or dense impactors partially dispersed by aerodynamic breakup might also produce central peaks by this mechanism. Too little is known of the appropriate conditions to be certain at present.

Finally, a third type of central peak has been observed to form in small-scale explosion craters with, again, very shallow depths of burial (Piekutowski, 1977). In these craters gases from the explosive charge scour the adjacent surface, leaving a central mound as an erosional remnant. Central peaks in some small experimental craters produced by highly dispersed impactors (basically shotgun blasts) also appear to be erosional remnants. Since such central peaks have zero or negative stratigraphic uplift, neither gas erosion nor a dispersed impactor can be the explanation of central peak formation in most complex craters.

### 8.3.3 Collapse to complex craters

The geologic and morphologic evidence reviewed in the previous two sections strongly supports the idea that a complex crater develops from a bowl-shaped transient crater by a process of gravity-driven collapse. The floor underlying the transient crater rapidly rises a distance comparable to the initial crater depth. Since breccia lenses are not observed in the central region of these craters, the collapse must be so rapid that debris has no chance to slide down the walls of the transient

crater. The timescale for the rise of central peaks is thus no longer than  $(D/g)^{1/2}$ .

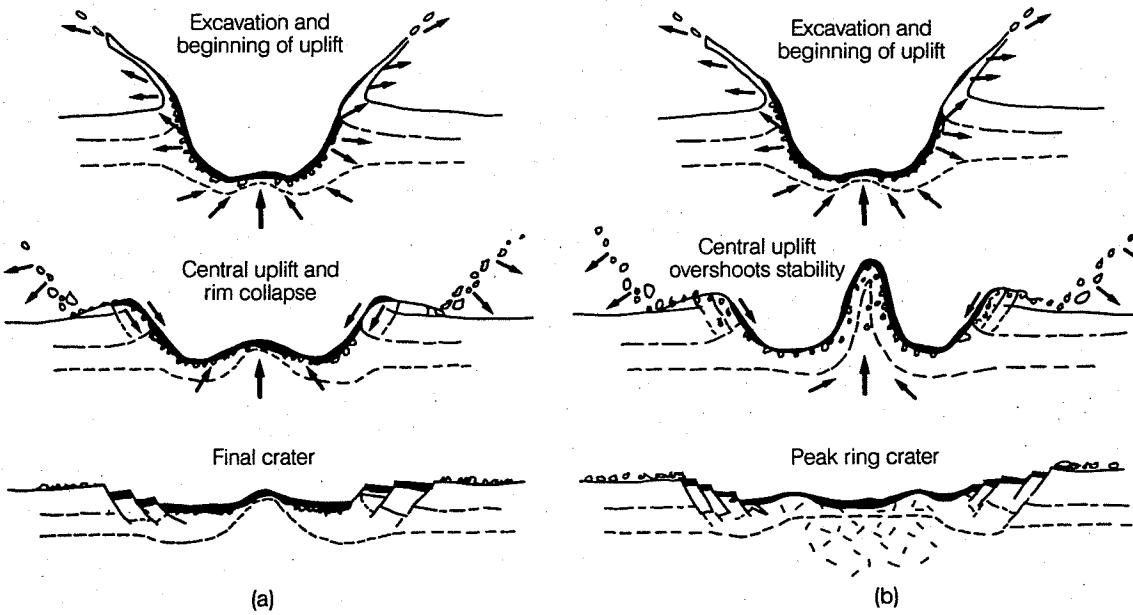
The breccia in complex craters fills an annular moat between the central uplift and the crater rim. Some breccia and melt may also be draped over the central peaks themselves. The floors of these craters are veneered with breccia and meltrock lying in the same stratigraphic sequence with which they lined the transient crater cavity. This stratigraphy explains the rather rough, crackly appearance of the floor of Copernicus and other fresh complex craters as the result of rapid cooling and shrinkage of the surface melt layer, which may also partially drain into the cooler, less shocked breccia underlying it.

Figure 8.14a schematically illustrates the formation of a complex crater and its central peak or peak ring. Most of the details of this figure are derived directly from the geologic study of terrestrial complex craters. One interpretative feature included in this illustration, however, is the depiction of the floor uplift beginning *before* the rim has been completely excavated. A fully formed parabolic transient crater thus never develops because the floor begins to rise almost as soon as the crater has stopped growing in depth. This type of collapse has been observed in laboratory-scale impacts into strengthless materials (liquid water) and has appeared in recent numerical simulations of the excavation of large impact craters.

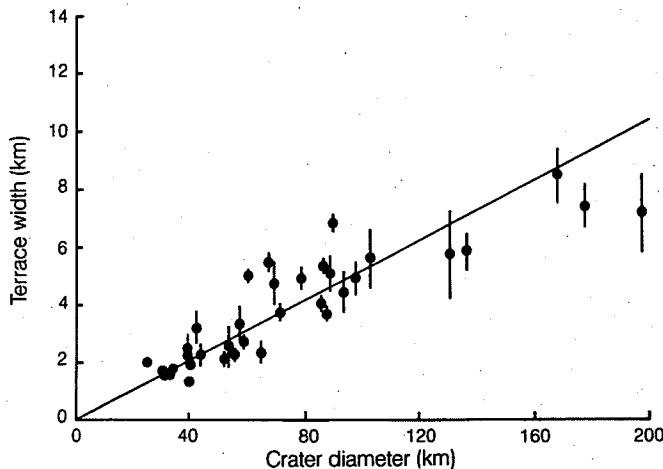
The wreath of terraces surrounding the crater floor must also develop very quickly after the impact. The toes of these terraces fade smoothly into

the solidified impact melt covering the crater floor without any sign of disruption by movement after the melt solidified. The terraces themselves are frequently veneered with impact melt that ponds in closed depressions formed by the backward-rotated terrace headscarsps. Leveed channels and flows lead from higher terraces to lower ones or onto the crater floor. These observations prove that the terraces were formed before the impact melt solidified. Since some of this "melt" is probably Suevite—a mixture of superheated melted rock and cold clasts that solidifies almost as soon as it is emplaced—the terraces probably formed as soon as the rim emerged from the excavation flow.

The terraces appear to be the headscarsps of great landslides that carried large amounts of material from the unstable crater rim into the interior of the crater. They resemble the "Toreva blocks" of classical geomorphology that are created by large landslides in North America's Colorado Plateau country. Terraces within an individual crater are generally narrow toward the crater interior, but increase in width as the rim is approached. The widest and best-defined terrace normally occurs just below the crater rim. The width of this largest terrace increases as the crater diameter increases, ranging from 1-km wide in 20- to 40-km diameter lunar craters to 8 km wide in craters between 160 and 200 km in diameter (Fig. 8.15). The kilometers-high scarp between the rim and this terrace also becomes higher in proportion to crater size. This scarp is very steep in



**Fig. 8.14** Schematic illustration of the formation of complex craters with either (a) central peaks or (b) peak rings. Uplift of the crater floor begins even before the rim is fully formed. As the floor rises further, rim collapse creates a wreath of terraces surrounding the crater. In smaller craters the central uplift "freezes" to form a central peak. In larger craters the central peak collapses and creates a peak ring before motion ceases.



**Fig. 8.15** The width of the largest terrace in complex lunar craters as a function of crater's diameter. The largest terrace forms just under the crater rim and is the last to slip into the crater's interior. It is also the best defined of all the terraces. In general, terrace width increases from the inside of the crater outward, so that earlier-formed terraces are smaller than late ones. After Pearce and Melosh (1986).

fresh lunar craters, commonly standing at slopes of 30 degrees, near the angle of repose.

One of the mechanical mysteries of crater collapse, others of which we shall consider in the next few sections, is that terraced features are typical of the headscarsps of landslides that form in *plastic* materials such as water-saturated clays. Toreva blocks form where massive sedimentary rock units are underlain by weak plastic clay layers. However, there is no water on the moon. In the absence of water it is expected that lunar slope failure should resemble that of dry sand or other materials dominated by Coulomb friction. Crater rims should fail by progressive sloughing of surface layers until the slope has collapsed back to the angle of repose. No terraces at all should form. This does, indeed, describe the sliding of debris in simple craters, but it is a poor description of failure in complex craters. The average slope of the terrace zone is between 7 and 14 degrees for complex craters in the 30- to 200-km diameter range, far below the normal angle of repose of loose rock debris. The resolution of this mystery is controversial, but it clearly involves transient mechanical effects associated with crater excavation.

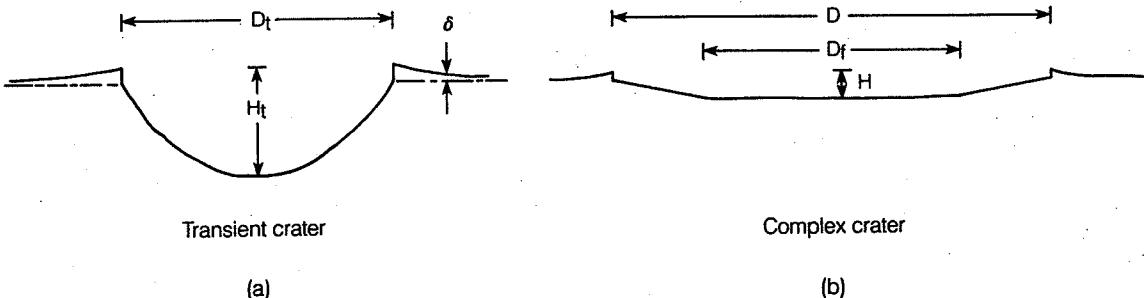
Peak-ring craters are thought to form when the central uplift itself collapses after an initial rise (see Fig. 8.14b). This is borne out by morphometric data (Sec. 8.3.1), which shows that central peak height stops increasing in lunar craters larger than about 80 km in diameter. Peak rings are faintly discernible in larger craters and completely replace central peaks at diameters larger than 140 km. The motion of rock debris in the crater interior must thus be very fluidlike, involving rapid uplift of a central peak, analogous to the central jet that forms when a cavity in water collapses. When the central peak collapses, a ring of mountains surrounding its former site rises, analogous

to the first ripple that develops around a rock thrown into a pond. The major difference between flow in a crater and that in a fluid is that the motion in a crater is arrested or "frozen" in some intermediate state that depends upon the crater size. How this happens, and what causes the apparent fluidity of rock debris, is another of the mechanical mysteries associated with crater collapse that we shall deal with in subsequent sections. For now it is sufficient to cite the kinematic truth that this motion *does* seem to happen, even if the explanation of *how* it happens escapes us for the moment.

Before delving further into the mechanics of crater collapse, the dimensions of the transient crater will be deduced from measurements on the final crater using a geometric model similar to that used for simple craters in Section 8.2. The essential features of this model are shown in Figure 8.16. The observed complex crater is characterized by its rim-to-rim diameter  $D$ , its rim-to-floor depth  $H$ , and the floor diameter  $D_f$ . The width of the terrace zone is thus  $D - D_f$  and its average slope is  $(D - D_f)/H$ . The volume of the central peak complex is generally much less than the crater volume and may be neglected in the volume balance. The parabolic rimmed transient crater of initial diameter  $D_i$  and depth  $H_i$  is assumed to collapse in such a way that the volume of the solid material surrounding it is conserved. If the rim height  $\delta = H_i/40(D_i/r)^3$ , so that the rim volume equals the crater volume, a simple relation exists between the theoretical quantity  $D_i^3 H_i$  and observable quantities:

$$D_i^3 H_i = \frac{5D^3 H}{1 + 6\{1 + (D_f/D) + (D_f/D)^2\}^{-1}} \quad (8.3.1)$$

## IMPACT CRATERING



**Fig. 8.16** Geometric model of complex crater collapse. Complex craters' small depth/diameter ratio compared with simple craters means that the final crater diameter  $D$  is much larger than the transient crater diameter  $D_t$ .

If an independent relation between the depth and diameter of the transient crater can be obtained, this equation yields the diameter (or depth) of the transient crater that collapsed to form a given complex crater. It does not matter whether the transient crater ever completely forms as a parabolic crater, so long as the thickness of primary ejecta  $\delta$  is the same whether immediate collapse occurs or not. The transient crater depth and diameter obtained from Equation 8.3.1 are still correct, although they may never be achieved *simultaneously* by the transient crater. The formula used for the rim height also neglects incorporation of preexisting surface material in the ejecta blanket by ballistic sedimentation (see Sec. 6.3.3). This is actually correct, since reworked surface debris does not add to the volume of material sliding back into the crater: only the primary ejecta that builds the rim up above the preexisting surface does this. Equation 8.3.1 does, however, ignore the increase in ejecta volume that is due to brecciation.

If the depth/diameter ratio of the transient craters that form large complex craters is assumed to be about 1/2.7, the same as for simple craters (Sec. 8.2), then Equation 8.3.1 may be solved to obtain the transient crater diameter and thus estimate the total expansion of the crater rim during collapse. For example, a 100-km diameter lunar crater such as Copernicus or Theophilus has a floor diameter of 60 km and a depth of 4.2 km (see Fig. 8.7). In this case Equation 8.3.1 predicts a transient crater diameter of 62 km, nearly equal to the floor diameter. The transient crater thus expands by 60 percent of its initial diameter during the course of crater collapse. Many authors have found similar values for the expansion of the transient crater, so this result may be accepted as typical.

The rim height  $h_R$  of the final crater can also be estimated by an elegant extension of Equation 8.3.1. This height is given by

$$h_R = \frac{H}{1 + 6(1 + (D_f/D) + (D_f/D)^2)^{-1}} \quad (8.3.2)$$

The predicted rim height is independent of the assumed transient crater depth/diameter ratio. When the rim height computed from Equation 8.3.2 is compared with actual complex crater rim heights (Fig. 8.7 and Table 8.1), the predicted rim height increases with crater diameter at the same rate as the actual rim height, but the predicted height is about 40 percent too small. This discrepancy is probably due to incorporation of preexisting surface material into the ejecta blanket. Although such material should not be included in the volume balance that leads to Equations 8.3.1 and 8.3.2, it does increase the measured rim height.

### 8.3.4 Mechanics of crater collapse—phenomenology

The observed facts of complex crater collapse can be used to infer the mechanical properties of the material surrounding the crater from a few simple arguments. These properties are unlike any that have ever been suggested for rock debris. Nevertheless, if it is assumed that these properties really do describe the behavior of rock surrounding the site of an impact, a surprisingly accurate and comprehensive theory of crater collapse results. This type of model is thus purely phenomenological—it works, but *why* it works is not apparent. The next section will examine a physical mechanism that seems to explain why rock debris surrounding an impact crater behaves in such a peculiar fashion. Other unusual geologic phenomena, such as long runout landslides, appear to operate by a related mechanism.

The abrupt onset of complex crater collapse suggests that some definite strength threshold is exceeded beneath craters larger than a critical size. The size of this strength threshold can be roughly estimated by dividing the negative buoyancy force associated with the crater cavity,  $\pi/8 \rho g H D^2$ , by the area of a hemisphere enclosing the crater,  $\pi/2 D^2$ . The resulting mean stress is  $1/4 \rho g H$ . For a transient lunar crater near the threshold diameter of 15 km this stress is on the order

of 5 MPa—far lower than the average crushing strength of most rocks, which is typically 100 MPa.

A more precise (although not necessarily more accurate) estimate of the strength threshold is derived from the full solution for the stresses in an elastic half space containing a rimmed crater. The result of such an analysis (Melosh, 1976) is that the maximum shear stress occurs under the center of the crater at a depth of  $0.23D$  below the preimpact surface. The size of this maximum stress is  $0.36\rho gH$ —only slightly higher than the previous rough estimate.

It is clear that the stresses associated with collapse of craters are far lower than the measured strengths of intact rocks. However, the rocks beneath an impact crater are unlikely to be intact—perhaps an analysis employing the strength characteristics of loose rock debris would yield a better result. Loose rock debris, however, fails by a Coulomb friction law, for which sliding begins under applied shear stresses larger than some constant  $\mu$  (the coefficient of friction) times the overburden. Since geologic evidence clearly indicates that sliding takes place at depths at least as great as the crater depth  $H$  below the crater floor, Coulomb failure requires that shear stresses exceed  $\mu \rho gH$ . But  $\mu$  is typically 0.85 for rock debris of all types, so shear stresses must exceed 0.85  $\rho gH$ . Since the maximum stress estimated above is less than 0.36  $\rho gH$ , even the strength characteristics of loose rock debris cannot explain impact crater collapse. A more direct way of obtaining the same result is to note that the average internal slope of large complex craters is between 2 and 10 degrees, far lower than the 30-degree average angle of repose of rock debris, so that frictional sliding cannot account for the wholesale collapse that creates complex craters.

Since conventional rock and debris strength properties cannot explain crater collapse, the problem can be turned around to determine what strength properties the rock must possess to explain the observed collapse. Terrace morphology provides a clue to which failure law must be used: terraces are characteristic of the failure of *plastic* materials. Such materials fail when the shear stress exceeds a well-defined yield stress called the cohesion,  $c$ . Perfectly plastic materials lack internal friction: the cohesion is independent of overburden pressure. This failure law provides a fair description of metals or of water-saturated clays. It is *not* a good description of the strength properties of either intact rock or broken rock debris.

Standard soil mechanics techniques exist for evaluating the stability of structures in a plastic material against collapse driven by gravity. These techniques are successfully used for predicting the strength of engineering structures such as dams and railway embankments. Although the most

common method, slip line analysis, does not include inertial forces, it does give a good description of the actual failure surfaces that develop in the collapsing material. Inertial forces are likely to be important in the central regions of collapsing craters, where the velocity of inflowing material is high, but are probably negligible near the rim where the terraces form.

Slip-line analysis applied to the collapse of impact craters (Melosh, 1977) gives a very accurate description of their collapse. Parabolic craters are stable until the dimensionless parameter  $\rho gH/c$  exceeds 5. Slope failure, in which a segment of the rim slides into the crater, producing a single terrace, occurs for values of  $\rho gH/c$  between 5 and 10. When  $\rho gH/c$  exceeds 15, the floor beneath the center of the crater rises almost vertically upward as the rim slumps downward. In terms of transient crater diameter  $D_t = 0.27H$ , these criteria are

$$0 \leq D_t \leq 13.5 c/\rho g \dots \dots \dots \text{Stable} \quad (8.3.3a)$$

$$13.5 c/\rho g \leq D_t \leq 27 c/\rho g \dots \dots \dots \text{Slope Failure} \quad (8.3.3b)$$

$$27 c/\rho g \leq D_t \leq 40 c/\rho g \dots \dots \dots \text{Floor Failure} \quad (8.3.3c)$$

Note that for constant cohesion  $c$  and density  $\rho$  the transition diameters are inversely proportional to surface gravity,  $g$ , in agreement with the data plotted in Figure 8.5.

The theoretical stability predictions summarized in Equations 8.3.3a–8.3.3c have been verified by a series of laboratory experiments in which structures very similar to those observed in complex craters are produced by collapsing craters in clay or plasticine using a centrifuge (see Fig. 8.17). An extension of the theoretical work (McKinnon, 1978) has shown that the material surrounding the crater must have a coefficient of friction less than about 0.035 because the observed stratigraphic uplift of the material underlying the crater center is suppressed for larger coefficients of friction.

Since the failure of purely plastic material is independent of overburden pressure, the collapse threshold is independent of the initial crater's diameter—only its depth is significant. As the crater collapses, material beneath it flows until the stress differences beneath all portions of the crater floor drop below the yield criterion, at which point collapse ceases. This model thus predicts that the final depth  $H$  of all complex craters is a constant equal to the depth of a transient crater at the onset of collapse—about 4.5 km on the moon, taking into account the thickness of the breccia lens in simple craters just below the collapse threshold.

$$H \approx 5c/\rho g = H_{\text{threshold}} \quad (8.3.4)$$

(a)



(b)



**Fig. 8.17** An experimental investigation of crater collapse in a plastic material. A 2cm-diameter crater formed in plasticine was centrifuged at 500 g and the recovered crater sectioned. The section reveals a well-developed central uplift and terraces defined by deep-seated normal faults, visible in the section as offsets of the layers. No central peak developed, however. Photos courtesy D. E. Gault, who performed this unpublished experiment.

The fact that craters in a plastic material collapse back to a constant depth also accounts for the development of a broad flat floor in the interior of large complex craters. Figure 8.4 shows that the depth of complex lunar craters does depend somewhat upon diameter, but it only varies between about 2.5 and 6 km over a diameter range of 20 to 300 km. This variation could be explained by a very small coefficient of friction or by a small dependence of the effective cohesion on crater diameter.

The plastic collapse model predicts the width of each individual terrace as it forms. A simple analytic formula that gives a good approximation to the numerical slip line results for the terrace width  $w$  is

$$w = \frac{c}{\rho g} \left( \frac{1 + 16\lambda^2}{16\lambda^2} \right) \quad (8.3.5)$$

where  $\lambda$  is the depth/diameter ratio of the crater at the time when the terrace forms. The first terrace to form, when the depth/diameter ratio equals that of the transient crater, is thus the narrowest, whereas the last terrace to form, when the crater's depth/diameter ratio is only slightly larger than its final value, is the largest, as is qualitatively observed in individual craters.

A quantitative test of Equation 8.3.5 within an individual crater is difficult to perform because the crater's depth/diameter ratio at the time a particular terrace formed is nearly impossible to determine. It is easier to compare the widths of the largest and last terrace to form (Fig. 8.15) between different craters, since it is possible to infer the depth/diameter ratio just before the final collapse. Figure 8.18 compares the terrace width predicted

by Equation 8.3.5 for a number of different values of cohesion to the observed terrace widths in a number of lunar craters. The agreement with this simple formula is impressively good and implies cohesive strengths in the neighborhood of 2 MPa. This strength also agrees reasonably well with the value of 3 MPa independently derived from the crater depth at the onset of collapse (see Table 8.2).

These results strongly support the idea that the strength properties of the material beneath a collapsing crater are correctly described by a perfectly plastic material with negligible internal friction and a cohesion of about 3 MPa. The only major feature that is not correctly described (and did not appear in a sequence of experimental runs with plastic materials in a centrifuge) is the central peak complex. Some other material property is evidently necessary to explain the formation of central peaks.

Central peaks have frequently been likened to the central jet that rises out of a collapsing cavity in water (Fig. 8.19). If central peaks in impact craters actually do form by a hydrodynamic mechanism of this kind, the rock debris beneath the crater must behave as a fluid for at least a short period of time. An upper limit on the effective viscosity  $\eta$  of this fluid can be derived from the 1890 work of P. G. Tait. Tait found that a crater in a viscous fluid collapses without overshoot if the viscosity is greater than

$$\eta_{\text{crit}} = \pi^{3/2} \rho g^{1/2} D^{3/2} \quad (8.3.6)$$

If the viscosity is less than  $\eta_{\text{crit}}$ , a central peak forms. This peak may oscillate one or more times before its motion is damped out, depending on the viscosity. The central peak is, in effect, a damped harmonic oscillator. If the damping is larger than critical (that is, the viscosity is larger than  $\eta_{\text{crit}}$ ), then the crater cavity simply collapses without oscillation. If the damping is less than critical (viscosity less than  $\eta_{\text{crit}}$ ) then the crater center executes a number of up-and-down oscillations of exponentially decreasing amplitude before it comes to rest.

Inserting parameters appropriate for a 20-km diameter lunar crater into Equation 8.3.6, one finds that the debris surrounding a crater of this size must have a viscosity less than about  $6 \times 10^{10}$  Pa-second for central peaks to arise hydrodynamically. As crater size increases a hydrodynamic central peak would first rise then collapse to form a circular ring. Repeated oscillations might be expected in even larger craters.

The collapse of impact craters cannot be strictly hydrodynamic, however, for the end result of the collapse of a crater in a viscous fluid is a plane surface without any topographic relief. Evidently, the hydrodynamically rising jet must be "frozen" or somehow halted to produce the observed crater forms. This explanation of complex crater mor-

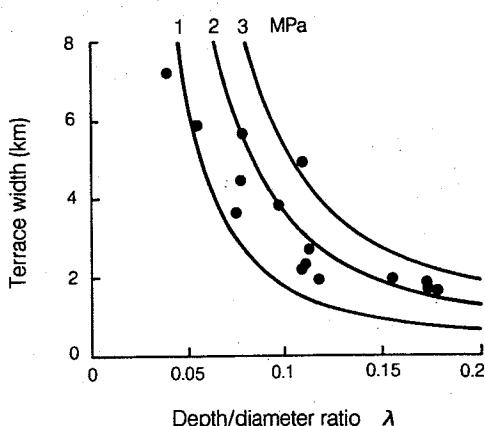
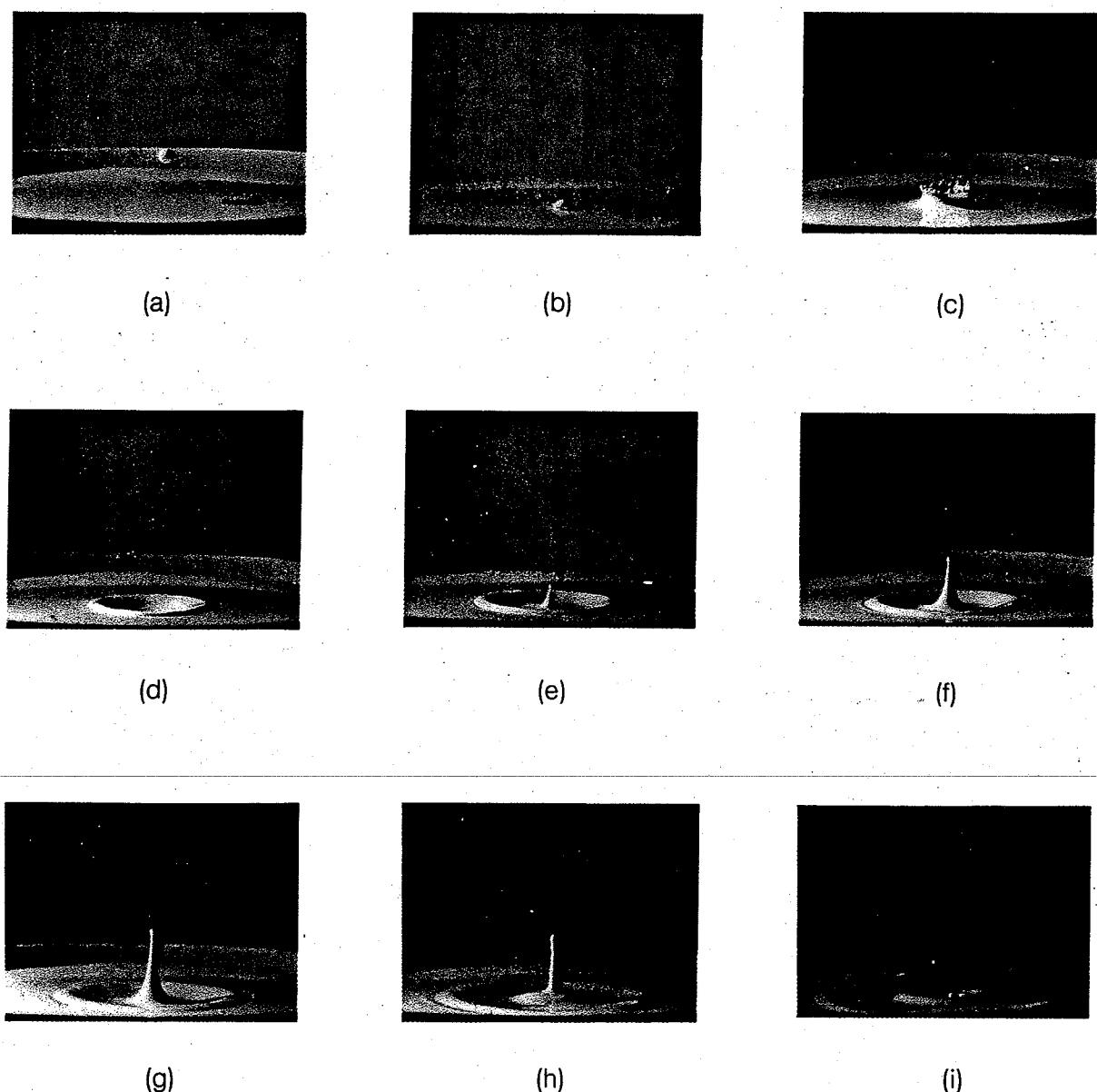


Fig. 8.18 Lunar crater terrace widths (dots) from Figure 8.15 are plotted against the depth/diameter ratio inferred for the crater just before the final terrace formed. This data is well expressed by the theoretical curves from Equation 8.3.5 assuming a cohesive strength between 1 and 3 MPa. These results tend to support the plastic slump model. After Pearce and Melosh (1986).

## IMPACT CRATERING



**Fig. 8.19** The hydrodynamic theory of central peak formation is illustrated by the central jet raised in water following the impact of a droplet. The central jet first rises, then collapses, forming a ring around its former location. *Photo sequence courtesy R. B. Baldwin (Photographs by Gene Wentworth of Honeywell Photograph Products).*

phology was first proposed privately by R. B. Baldwin about 1966. Baldwin inspired oceanographer W. G. Van Dorn (1968, 1969) to undertake detailed computations using the theory of long-wavelength ocean waves, or tsunami. The resulting "frozen tsunami" theory (Baldwin, 1972, 1981) did seem to explain the origin of central peaks and peak rings in complex craters qualitatively, but many planetary scientists objected to the ad hoc introduction of the "freeze" and could not accept what appeared to be a mechanically inconsistent theory.

On the other hand, it appears that many other aspects of complex crater collapse can be de-

scribed as if the material surrounding the crater is perfectly plastic. Although it may seem as though plastic and viscous fluid properties are contradictory, there is a rheological model that incorporates both types of behavior, each being expressed in its own domain of validity. This model not only reconciles the now-plastic, now-fluid behavior of material surrounding the crater, but it also provides a mechanical basis for the "freeze" and permits quantitative analysis of the process of crater collapse.

A material that behaves both as a viscous fluid and has a definite plastic yield stress is known as *Bingham fluid*. This rheologic model was first in-

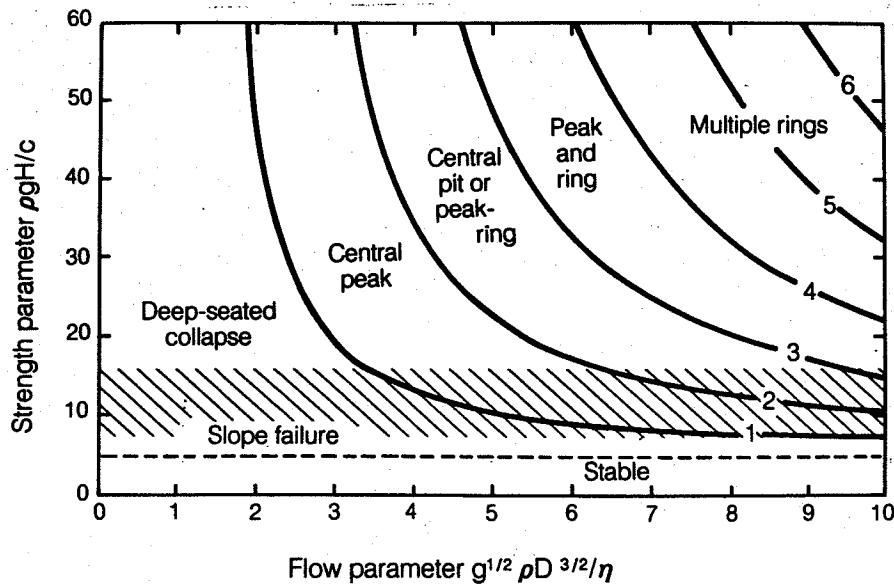
roduced in 1916 to describe the behavior of clay suspensions and paint. A Bingham fluid supports applied stresses without appreciable strain until the shear stress exceeds its cohesion  $c$ , after which it flows as a fluid with viscosity  $\eta$ .

A central peak rising out of a crater formed in a Bingham fluid continues to move until the shear stresses driving it fall below the cohesion, at which point its motion ceases and it appears to "freeze." Even though the central jet may briefly rise many times the transient crater depth, as in Figure 8.19 (or plunge some distance below it), the final central peak (or depression) in such a material cannot be higher (or much lower) than the depth of the final crater, otherwise shear stresses below the peak would exceed the cohesion. Hydrodynamic flow continues until all stresses fall below the cohesion. This prediction is roughly borne out by the data in Figure 8.7, in which it can be seen that central peak heights seldom exceed the crater depth. A phenomenological model of crater collapse based on the properties of a Bingham fluid thus appears to satisfy most of the observations and inferences on the collapse of complex craters on the terrestrial planets. Experimental impacts into Bingham fluid targets (clay slurries) also yield crater forms similar to those observed in complex craters (Greeley et al., 1980; Fink et al., 1981). The major facts that have not yet been explained on the basis of the Bingham (or any other) model are the apparently universal ratios of central peak diameter to final crater diameter (roughly 0.22) and the ratio of

peak-ring diameter to final crater diameter (roughly 0.5). It may be that these ratios actually do result from the Bingham model: the dynamical computations to find out have not yet been performed. These ratios, however, are suggestively close to the ratios between some of the zeros of the  $J_0$  Bessel function, which is expected to be important in describing the sloshing motion of a fluid confined in a circular cavity.

A detailed model of crater collapse based on a Bingham fluid rheology has been constructed (Melosh, 1982a). Figure 8.20 shows its principal predictions on a plot of dimensionless yield parameter,  $\rho g H/c$ , and flow parameter,  $\rho g^{1/2} D^{3/2}/\eta$ , which is essentially the ratio between  $\eta_{cri}$ , Equation 8.3.6, and viscosity  $\eta$ . The major internal features such as central peaks, peak rings, and possible multiple oscillations of the fluid inside the crater appear on this plot. Lower-speed collapse near the crater rim produces the terrace zone at the same time as these hydrodynamic features form in the crater center.

The plot in Figure 8.20 is not useful for predictive purposes, however, unless some values are available for the phenomenological Bingham parameters, the cohesion  $c$  and viscosity  $\eta$ . These values are deduced from the crater diameter at which a number of important transitions occur on various planets and satellites in Table 8.2. This table shows that the cohesive strength  $c$  is similar on the terrestrial planets, although the cohesion on Mars is somewhat smaller than on the other planets. This result merely reflects the  $1/g$  depen-



**Fig. 8.20** The phenomenological model of crater collapse predicts that the crater's final form depends upon two dimensionless combinations of the Bingham parameters cohesion  $c$  and viscosity  $\eta$  with the target density  $\rho$ , surface gravity  $g$ , crater depth  $H$  and diameter  $D$ . For small values of the flow parameter, the crater collapses in a purely plastic manner in which no central peak rises. At larger values of the flow parameter, central peaks or rings may form, provided the strength parameter is sufficiently large. After Melosh (1982a).

## IMPACT CRATERING

TABLE 8.2 *Bingham parameters for crater collapse*

Planet or Satellite	Crater Diameter (km) at Onset of:			Cohesion* <i>c</i> MPa	Viscosity† <i>η</i> GPa-sec
	Collapse	Central Peaks	Peak Rings		
Mercury	4.8‡	<10**	—	2.8	<5.2
Earth	1.9‡	2.5**	75**	—	53.0
Moon	12‡	<10**	—	3.0	1.1
Mars	3.1‡	3**	—	3.2	<3.4
Ganymede	—	—	140**	—	90.0
Rhea	—	5**	—	—	0.42
Dione	—	15††	—	—	1.0
Tethys	—	20††	—	—	1.4
Enceladus	—	27††	—	—	2.2
Mimas	—	15††	—	—	0.55
	—	27††	—	—	1.3

\*Determined from  $c = 0.06 \rho g D$ , where  $\rho = 2700 \text{ kg/m}^3$  for silicates,  $\rho = 1000 \text{ kg/m}^3$  for icy bodies.

†Determined from  $\eta = \rho \sqrt{g D^3}$  for central peaks, half this value for peak rings.

‡Pike (1988)

\*\*Pike (1983)

††Chapman and McKinnon (1986)

dence of the simple-to-complex transition diameter illustrated in Figure 8.5 and the plastic collapse model's prediction that transition diameter scales as  $1/g$  for constant cohesion. The cohesion on Ganymede is significantly less than that on the terrestrial planets. Cohesion on the other icy satellites in the outer solar system cannot be estimated because the images presently in hand lack sufficient resolution to determine the simple-to-complex transition.

The viscosity at the onset of large central peaks is variable, but within broad limits is roughly similar among both the terrestrial planets and the icy satellites. The viscosity at the onset of peak-ring formation, however, is 30 to 50 times higher on the moon than on the earth or Mars. Moreover, on each planet it is nearly 30 times larger than the viscosity of the onset of central-peak formation. This suggests that the Bingham viscosity increases as crater size increases. Terrace width measurements over the same range of crater diameters provide no evidence that cohesion increases markedly with diameter, however. The understanding of these relations awaits a complete physical theory of the mechanism of crater collapse.

The phenomenological Bingham fluid model gives a good quantitative description of the formation of complex craters as a result of gravity-driven collapse of an initial, deep transient crater. The Bingham parameters are not, however, comprehensible from any simple point of view and actually seem to be in conflict with the currently understood strength properties of rock debris. The value of the phenomenological model, besides providing an accurate description of complex cra-

ter formation, is to suggest areas of research that might elucidate the problem and that may result in a physical explanation of the mechanics of crater collapse. The next section discusses one such physical explanation of the peculiar parameters that underlie the Bingham fluid model.

Before going into this physical model, however, it is intriguing to note that a Bingham model rather similar to the one described earlier has been proposed to describe the flow of large rock avalanches that involve more than about  $10^5 \text{ m}^3$  of dry rock debris (Sturzstroms) (Hsu, 1975). These avalanches begin in a variety of ways, but once they have fallen some 1000 m, they continue to flow as if they had very little internal friction, moving down very shallow slopes or even occasionally coasting upslope. The Bingham viscosity (deduced from the thickness and flow velocity) of such landslides is about  $10^4 \text{ Pa-second}$  and the cohesion (deduced from the thickness of the final deposit) is 0.5 to 1 MPa, parameters that are similar to those deduced from crater collapse, and equally difficult to explain. Crater collapse and large dry landslides thus appear to have much in common, and the explanation of one is likely to apply in some way to the other.

### 8.3.5 Mechanics of crater collapse—physics

The phenomenological Bingham fluid model of crater collapse described in the preceding section provides a good quantitative description of the mechanics of impact crater collapse. The Bingham parameters, cohesion  $c$  and viscosity  $\eta$ , however, are determined by fitting the model's predictions to observations of collapsed craters.

The model gives no means of determining these parameters from the fundamental physical properties of the material in which the crater forms. Even worse, the Bingham parameters seem to disagree with traditional notions of the behavior of rock debris.

Several ideas have been advanced to explain the apparent low strength and fluid behavior of rock debris surrounding an impact crater. All of these ideas exploit the rapidity of the collapse to propose some transient mechanism that acts only briefly after the crater is excavated. Once crater collapse has occurred, the rock debris regains its familiar static strength characteristics, avoiding any fundamental paradox in the strength of rock materials.

One idea that was proposed early in the study of crater collapse is that the rock debris beneath the crater is fluidized by impact melt. The debris flows briefly as a melt-solid slurry until it cools and solidifies. Some impact melt is found on the floors of complex craters, but it is rare in the central uplifts or in the zone of stratigraphic uplift beneath the crater, where fluidization is required. The volume of impact melt surrounding known complex craters falls far short of the volume required to fluidize much of the surrounding rock debris. Furthermore, it is difficult to see how the production of impact melt alone can account for the observed sudden onset of collapse in craters at the simple-complex transition. This mechanism is thus implausible for craters smaller than a few hundred km in diameter on the terrestrial planets, although it might be important in very large craters for which the volume of impact melt is comparable to the crater volume (this occurs for transient craters larger than about 2000-km diameter on the moon and 300-km diameter on the earth; see Sec. 7.10) or on icy satellites where the surface material is more easily melted than silicate rocks.

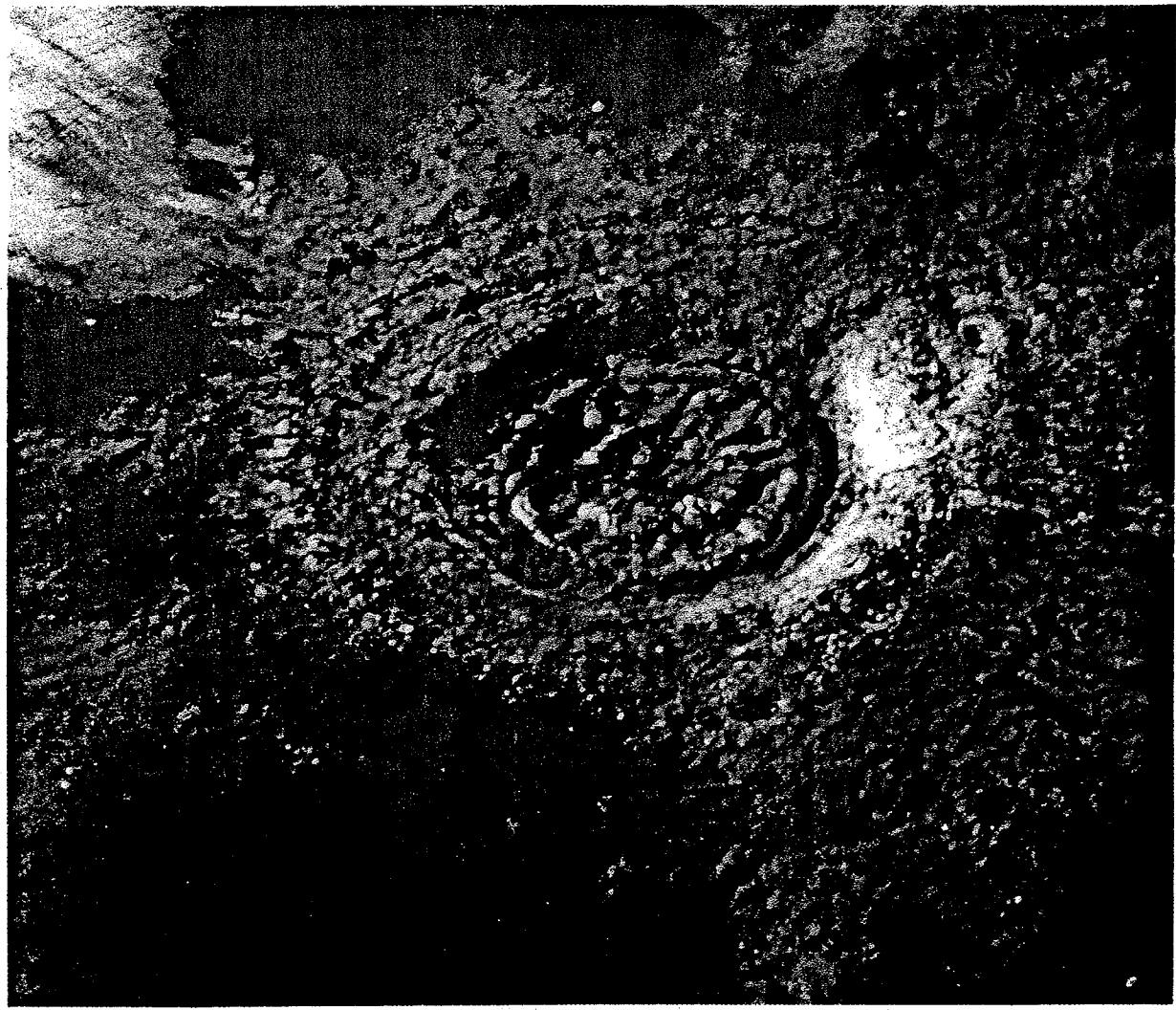
Another idea that is irrelevant for lunar or Mercurian craters but which may apply elsewhere is that the debris surrounding the crater is fluidized by the presence of some interstitial fluid, usually liquid water. Explosion craters on earth are strongly influenced by the presence of subsurface water. The widely cited 86-m diameter Prairie Flat Crater (Fig. 8.21), produced by the detonation of a 500-ton TNT sphere on water-saturated silts and clays in Canada, possesses multiple internal rings analogous to those observed in complex lunar craters. The shallow 700-m-diameter nuclear crater KOA on Eniwetok Atoll, which has a depth/diameter ratio of only 1/13, is now accounted for by the collapse of a much smaller diameter transient crater when the water-saturated coral rock beneath the crater liquified and flowed inward to partially fill the crater cavity. The exaggerated development of central peaks observed in many Martian craters may also be due to near-surface liquid water, whose presence is suggested

by the fluidized appearance of Martian ejecta blankets (Sec. 6.3.5).

It is well known that water-saturated silts and clays or slurries of rock debris and water behave like Bingham fluids. It is thus no surprise that craters produced in such materials on the earth resemble extraterrestrial complex craters (see also the experiments reported by Greeley et al., 1980, and Fink et al., 1981). However, the cohesion of these materials is typically  $10^{-4}$  to  $10^{-3}$  MPa, many orders of magnitude smaller than that inferred from the collapse of the extraterrestrial craters (Table 8.2). The viscosity observed in terrestrial mudflows is in the vicinity of  $10^2$  to  $10^3$  Pa-second, also many orders of magnitude smaller than observed in extraterrestrial crater collapse. This is, after all, why features characteristic of complex craters are observed in the much-smaller-scale impact and explosion cratering experiments. Furthermore, however applicable water-mediated fluidization may be to terrestrial or Martian craters, this mechanism is incapable of explaining the collapse of lunar or Mercurian craters. Even on earth it is difficult to argue that rock debris beneath large craters such as the Ries Crater was fluidized by water, since the original crystalline rocks possessed little initial porosity. Brecciation, in fact, should have greatly decreased the fluid pore pressure and prevented any vestige of fluidization by water.

I have proposed a third mechanism for fluidizing debris beneath an impact crater. This mechanism, although still somewhat speculative, operates in the absence of either air or water and yields Bingham parameters close to the values inferred from the collapse of extraterrestrial craters. It is a natural result of crater excavation processes and even enjoys some experimental support. Although the last word on the mechanism that permits craters to collapse has certainly not been written, I will end this section with a description of the process of *acoustic fluidization*.

The basic idea behind acoustic fluidization is that rock debris subject to strong vibrations can flow like a fluid even in the absence of air or water. The vibration is transmitted as a sound wave via rock-to-rock contacts. Because the pressure in a sound wave alternates between more-than-ambient in the compressive parts of the wave to less-than-ambient in the rarefactions, failure and sliding of the rock debris may occur sporadically wherever a particularly strong rarefaction cancels most of the overburden pressure. Random pressure excursions result in localized regions of failure occurring at a steady rate throughout the debris mass (Fig. 8.22). The net result is a creeping motion of the rock debris that produces a finite strain rate even for applied shear stresses far below the threshold of motion for a Coulomb material. The strain rate is negligibly slow for common amplitudes of vibration, but



**Fig. 8.21** The 86-m diameter Prairie Flat explosion crater was created in water-saturated silts and clays that fluidized after the explosion, forming a Bingham plastic medium surrounding the crater. A crater with several internal rings is the result. (Photo courtesy D. Roddy.)

when the mean pressure fluctuation in the stress wave approaches the overburden pressure, the strain rate may become large, exceeding 1/second.

The strain rate  $\dot{\epsilon}$  in acoustically fluidized debris is given by a rather complicated function (Mellosh, 1979) of applied shear stress  $\tau$ , overburden pressure  $\rho g z$ , where  $z$  is the depth below the surface, the coefficient of friction,  $\mu$ , and the variance of the pressure fluctuations  $\sigma$ . It also depends upon the density of the debris  $\rho$ , the velocity of sound in the debris  $\alpha$ , and the dominant wavelength of the acoustic field  $\Lambda$ . In terms of these variables the strain rate is

$$\dot{\epsilon} = \dot{\epsilon}_y T [2/\text{erfc}[(1 - T)/\Sigma] - 1]^{-1} \quad (8.3.7)$$

where erfc is the complementary error function, and  $T = \tau/\tau_{\text{static}}$  is a dimensionless measure of applied shear stress, equal to the actual applied shear stress  $\tau$  divided by the shear stress  $\tau_{\text{static}} = \mu \rho g z$  required to induce sliding in the absence of vibration.  $T$  is less than or equal to one. The di-

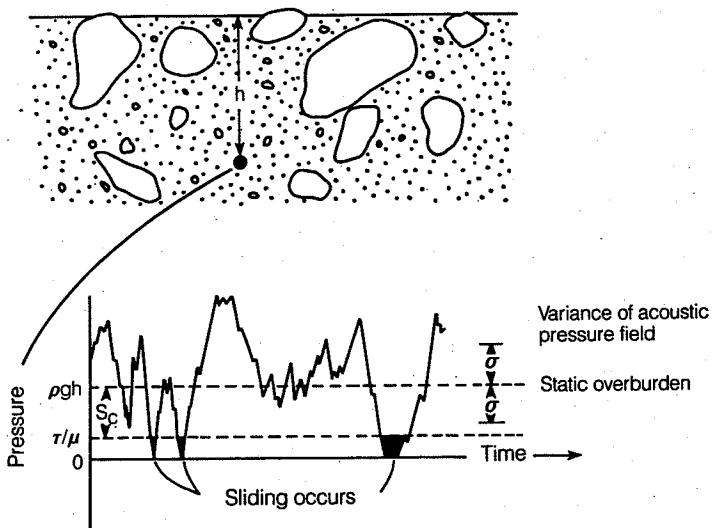
mensionless strain rate is  $\dot{\epsilon}_y = \tau_{\text{static}}/\rho \Lambda \alpha$ , and  $\Sigma$  is a dimensionless measure of the intensity of vibration,  $\Sigma = \sigma/\rho g z$ .

This rheologic law is plotted on a log-log scale in Figure 8.23, where it is also shown to agree with some experimental data. The figure makes it clear that when vibration is absent,  $\Sigma = 0$ , the strain rate is zero until the applied stress  $\tau$  reaches the static yield limit  $\tau_{\text{static}}$  (i.e.,  $T = 1$ ), at which point it is indeterminate from this model since vibration plays no role in conventional Coulomb failure. At intermediate vibration amplitudes, the strain rate is nonzero for applied stress below the static limit, although the strain rate may be exceedingly slow at very low stresses or for small amplitudes of vibration. In this intermediate regime the strain rate is proportional to a power of the stress larger than unity. Acoustically fluidized debris thus creeps according to a rheologic law similar to the one that governs the flow of hot crystalline rock or glacier ice. As the strength of

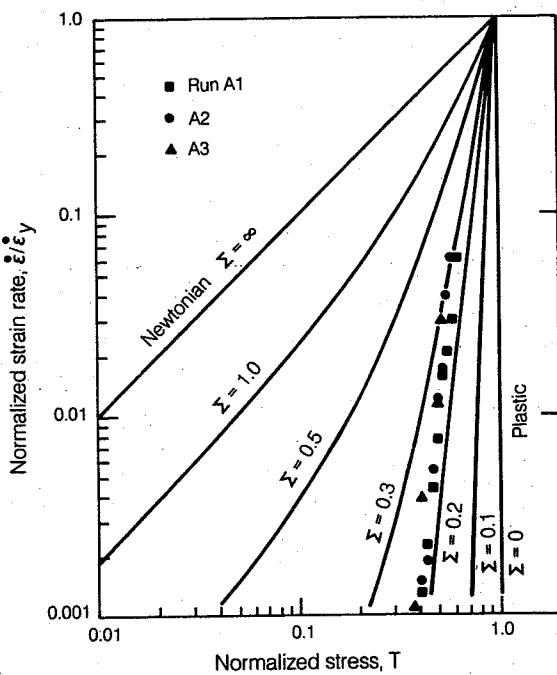
Fig  
po  
cas  
val  
roc

Normalized strain rate

Fi  
m  
th  
gi  
th  
ne  
T  
lc  
si  
pa



**Fig. 8.22** Dry rock debris may be fluidized by strong acoustic vibrations. The pressure at any given point at depth  $h$  oscillates about the static overburden  $\rho gh$  as sound waves pass by. The pressure occasionally drops below the level necessary for sliding to begin under an applied shear stress  $\tau$ . As the variance of the acoustic field  $\sigma$  increases, such sliding episodes become more frequent and the mass of rock debris deforms more rapidly.  $\mu$  is the coefficient of friction. After Melosh (1979).

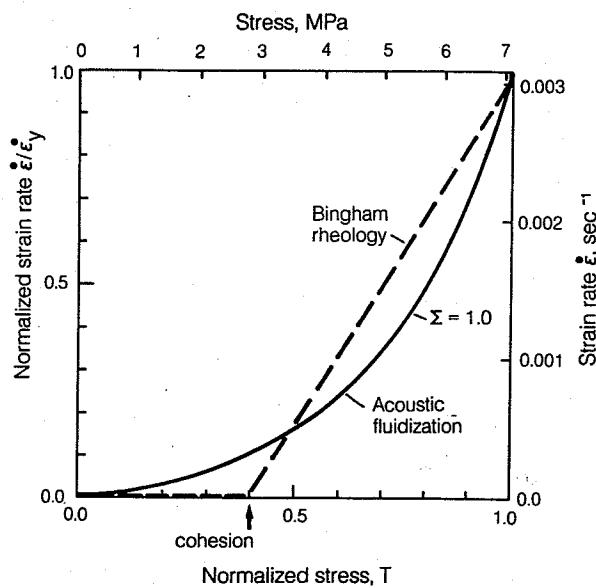


**Fig. 8.23** The strain rate of acoustically fluidized material is a function of the driving stress  $T$  and the variance of the acoustic pressure field  $\Sigma$ , here given in normalized form on a log-log plot. When the shear stress reaches the static limit ( $T = 1$ ) normal Coulomb-plastic deformation takes place. The presence of an acoustic field allows flow at lower stresses. Except for extremely strong acoustic field ( $\Sigma$  larger than about 10), the flow rate is strongly stress dependent. The plot shows a comparison between theoretical Equation 8.3.7 and a small quantity of experimental data. After Melosh and Gaffney (1983).

the acoustic field approaches infinity, the strain rate becomes directly proportional to the applied stress and the rheology of acoustically fluidized debris approximates that of a Newtonian fluid with viscosity  $\eta_{\text{eff}} = \rho \Lambda \alpha$ .

It is well known that the rheological behavior of a power-law fluid approximates a Bingham fluid. Figure 8.24 illustrates the relation between stress and strain rate for acoustically fluidized debris when the pressure fluctuations equal the overburden pressure ( $\Sigma = 1$ ) and a Bingham material. The strain rate is either low or zero in both materials until some stress (effectively a yield stress or cohesion) is reached, beyond which the material flows rapidly. An approximation of this type is commonly and successfully used to describe the flow of glacier ice, whose flow is actually governed by Glen's law,  $\dot{\epsilon} = (7.42 \times 10^{-5} \text{ sec}^{-1} \text{ MPa}^{-4.2}) \tau^{4.2}$ , but which is often represented as a Bingham fluid with a cohesion of 0.1 MPa and a viscosity of  $10^{13}$  Pa-second.

The top and right axes in Figure 8.24 are the stress and strain rate computed from Equation 8.3.7 for the collapse of a 20-km diameter crater on the moon. The static yield stress is taken to be 7 MPa, an average value down to depths of twice the transient crater depth. The acoustic field is assumed to have a dominant wavelength  $\Lambda \approx 900$  m, equal to the diameter of the 20-km/second projectile that excavates the 20-km diameter crater. The velocity of sound  $\alpha$  in the debris at a depth of a few km in the moon is taken to be 2 km/second. The Bingham parameters that best approximate this flow law are  $c = 3$  MPa and  $\eta = 1.3 \times 10^9$  Pa-second; values in good agreement with the empirically determined Bingham parameters (Table 8.2). The process of acoustic fluidiza-



**Fig. 8.24** An illustration of how acoustic fluidization's stress-dependent flow law can mimic a Bingham plastic rheology. The left and bottom scales are dimensionless, as in Figure 8.23 (but the plot is linear in this case, not logarithmic). The top and right scales are dimensional and are adjusted to values likely to be realized in the excavation of a 20-km diameter impact crater on the moon (see text). The Bingham parameters are close to those derived from the phenomenological collapse model.

tion thus appears to provide a reasonably good approximation to the rheology of material surrounding collapsing complex craters.

The source of the acoustic energy that fluidizes the debris surrounding a large crater is the stress wave generated by the impact. Although often represented theoretically as a smooth pulse that rises then dies away to zero, actual gauge records of the stress wave near explosions show that after the passage of the wave peak, a long-lasting train of apparently random stress wave "noise" continues to disturb the material surrounding the crater. The mean amplitude of this vibration easily exceeds the overburden pressure at distances less than a crater diameter from the crater's center (Gaffney and Melosh, 1982). The dominant wavelength,  $\Lambda$ , in the acoustic field is roughly equal to the diameter of the source, that is, the projectile's diameter.

Figure 8.25 illustrates how a single spherically expanding stress pulse is scattered by small random variations in the elastic moduli of an elastic material and shows how the hemisphere behind the expanding  $P$ -wave fills with a random mixture of both  $P$ - and  $S$ -wave energy. This process, occurring around the site of an impact crater, produces the high acoustic energy density necessary for acoustic fluidization of the material surround-

ing the crater and thus accounts for the simple-to-complex transition. During the interval when the acoustic energy is large, flow occurs at low differential stresses and the debris in the crater may even slosh back and forth like a liquid, producing central peaks or peak rings (Fig. 8.26). The acoustic energy density is sufficiently large to permit this flow only in a limited region surrounding the crater. More distant debris is effectively rigid at all times, so the flowing material surrounding the collapsing crater is confined in a roughly hemispherical bowl. After crater collapse is complete, the acoustic energy dissipates and the rock debris beneath the crater regains its normal strength properties.

This picture of crater collapse facilitated by acoustic fluidization is still evolving: more laboratory experiments on the rheology of acoustically fluidized debris are needed and more theoretical work on the details of acoustically fluidized flow beneath a crater must be performed. However, the process of acoustic fluidization appears to provide at least a qualitative physical explanation of the peculiar strength properties possessed by the material surrounding the site of a large impact crater and it thus provides a rationale for the morphologic form of complex craters.

#### 8.4 VISCOUS DEGRADATION OF CRATERS

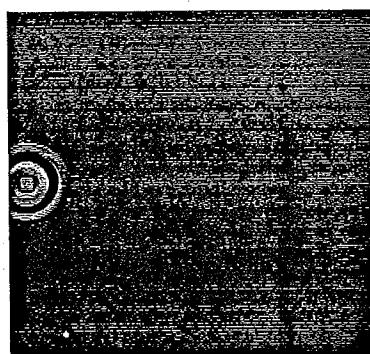
Simple and complex craters are the immediate result of impact crater excavation and collapse. However, long after these rapid events have occurred, craters on most planetary bodies continue to change and degrade toward the ultimate limit of gravitational stability: a level plain. One degradational process that seems to be particularly effective on the icy satellites of Jupiter (but which may also be important for large craters on Earth and other bodies) is the slow viscous flow of the material surrounding the crater.

No material is truly "solid." Research on the rheological behavior of crystalline solids during this century has shown that *all* substances flow or "creep" slowly under applied stress, however small the stress may be. The creep rate is strongly temperature dependent and, at ordinary temperatures, may be imperceptively slow, which leads to the common perception of many materials as solid. However, the creep rate  $\dot{\epsilon}$  of virtually all substances that have been studied attains a steady value some time after application of a differential stress  $\tau$  that is given by the equation

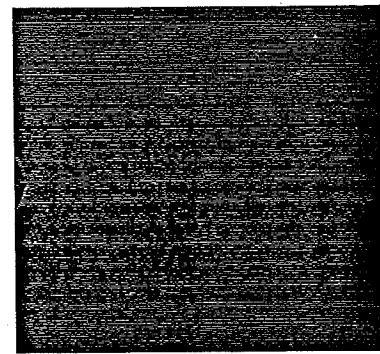
$$\dot{\epsilon} = A(\tau) \exp(-f T_m / T_A) \quad (8.4.1)$$

where  $A(\tau)$  is a function of the applied stress (in metals, ceramics, and geologic materials it is frequently of the form  $A\tau^n$ , where the power  $n$  is between 3 and 6),  $T_m$  is the melting temperature of the material,  $T_A$  is the absolute temperature, and

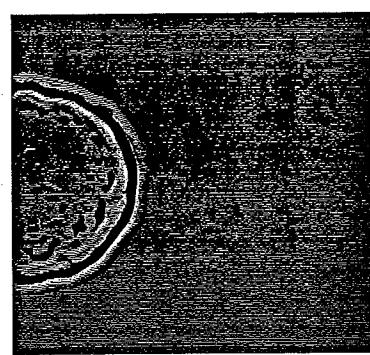
P-Wave (Divergence)



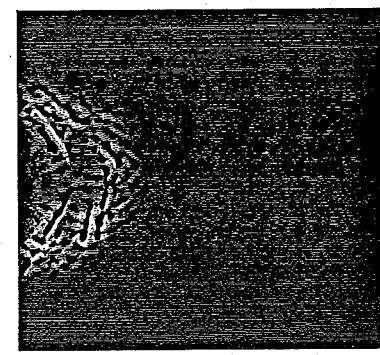
SV-Wave (Curl)



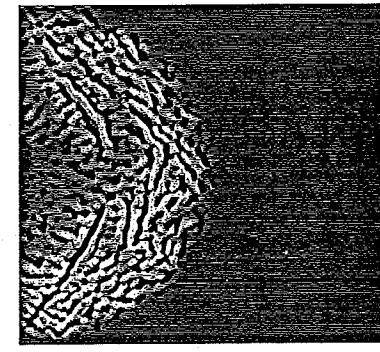
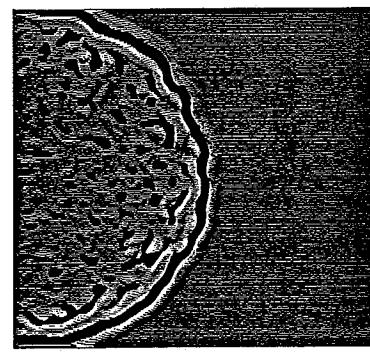
.064 sec



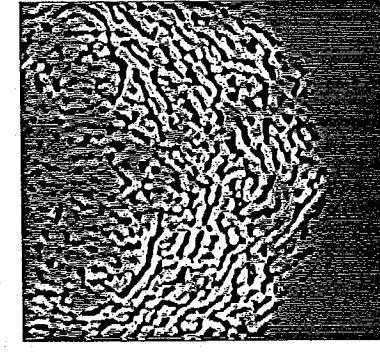
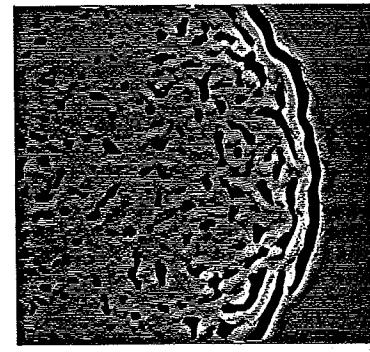
.192 sec



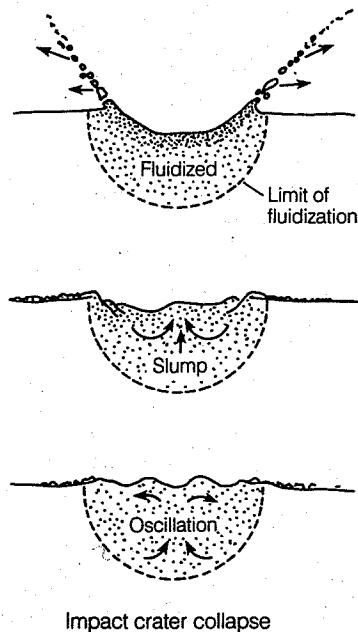
.320 sec



.512 sec



**Fig. 8.25** The generation of random acoustic energy from the propagation of a strong pulse in a randomly scattering medium. The pulse, originally a pure *P*-wave, quickly develops *S*-wave components by mode conversion in addition to leaving large amounts of *P* energy behind the advancing stress wave front. This illustrates how the random acoustic energy needed for acoustic fluidization may arise. *From a numerical experiment by Frankel and Clayton (1986).*



**Fig. 8.26** Schematic illustration of how acoustic fluidization participates in allowing the collapse and flow that forms complex craters. Because the acoustic field strength decreases rapidly with distance away from the crater, slumping and flow are strongly localized. *After Melosh (1983).*

$f$  is a dimensionless constant that is typically about 18 for metals and between about 25 and 27 for rock-forming minerals.

Although the creep rate in Equation 8.4.1 is negligibly small for rocks at the surface of the earth and over timescales comparable to a human lifetime, it becomes appreciable over geologic time intervals when the temperature approaches one-half of the melting point  $T_m$ . Thus, the base of the mechanical lithosphere in the earth's mantle is approximately defined by the 600°C (873 K) isotherm. At higher temperatures mantle rock behaves as a viscous fluid, permitting thermal convection and the slow drifting motion of the tectonic plates that tile the earth's surface. The flow rate of hot rock increases rapidly as the temperature approaches the melting point. The creep abilities of glacier ice, a crystalline substance very close to its melting point, are well known to all inhabitants of cold, snowy regions.

A convenient measure of the relative "solidity" of materials is the Maxwell time  $t_M$ . This is defined as the length of time a substance must flow under an applied load before the creep strain  $\epsilon t_M$  equals the elastic strain  $\epsilon_{el}$ . Thus, on timescales shorter than the Maxwell time a given substance may be regarded as an elastic solid, whereas as timescales longer than the Maxwell time the substance's fluid behavior predominates. In terms of the creep rate (Eq. 8.4.1), the Maxwell time is

$$t_M = \epsilon_{el}/\dot{\epsilon} = \eta/G \quad (8.4.2)$$

where the last expression  $\eta/G$  is the ratio between the substance's viscosity  $\eta$  and shear modulus  $G$ . The viscosity  $\eta$  is defined in terms of the strain rate (Eq. 8.4.1) by  $\eta = \tau/2\dot{\epsilon}$ . It is independent of stress only in the special case that  $A(\tau)$  is a linear function of  $\tau$ .

Upper crustal rocks of the earth and moon have viscosities in excess of  $10^{25}$  Pa-sec, and so their Maxwell times exceed  $10^9$  years—they behave as elastic solids essentially over the *full length* of geologic time. At the other extreme, the Maxwell time of glacier ice is only a few minutes, and ice's fluidlike properties are manifested on a timescale readily appreciated by humans. The cold ice on the surfaces of Ganymede and Callisto flows very slowly on a human timescale, but observations of craters on these large Jovian satellites show that appreciable flow has taken place over geologic time.

Because all materials are fluid over sufficiently long timescales, craters (or, indeed, any topographic feature) on the surface of any planetary object tend to flatten and disappear with time. The rate at which the depth  $H$  of a crater of diameter  $D$  decreases can be estimated from elementary considerations. The shear stress  $\tau$  beneath a crater of depth  $H$  is roughly  $\frac{1}{4} \rho g H$ , as derived in Section 8.3.4. The strain rate of material flowing inward toward the crater is thus about  $\dot{\epsilon} = \rho g H / 8\eta$ . The rate of uplift of the floor is roughly the crater diameter  $D$  times this strain rate,  $-dH(t)/dt = D\dot{\epsilon} = (\rho g D / 8\eta)H(t)$ . Solving this differential equation gives the crater depth as a function of time

$$H(t) = H(0) \exp(-t/t_R) \quad (8.4.3)$$

where  $H(0)$  is the initial depth and the relaxation time  $t_R$  is

$$t_R \approx 8\eta/\rho g D \quad (8.4.4)$$

There are a number of interesting features to note about this relaxation time. Most obviously, the relaxation time increases as the viscosity  $\eta$  increases. The relaxation time decreases as gravity or density increases. It also decreases as crater diameter  $D$  increases, so that large craters should relax faster than small ones. This is a general result of viscous relaxation on a deep layer of uniform viscosity: small-scale features persist longer than large ones. Even within a single crater the small-scale rim should persist longer than the broad crater bowl, a relation that is clearly observed within relaxed craters on Ganymede and Callisto.

Figure 8.27 illustrates a partially relaxed crater on Ganymede, with an unbowed floor and crisp rim, a more relaxed *penepalimpsest*, and a completely relaxed crater type called a *palimpsest*, whose existence is only betrayed by a bright albedo feature. All of these craters are inferred to have begun as complex crater forms that have re-



**Fig. 8.27** Palimpsest and penepalimpsest craters on Ganymede. The palimpsest at the lower left lacks both a central craterform depression and a rim. It is visible only as an albedo feature. The penepalimpsest crater in the center still possesses a recognizable rim. Both craters have relaxed by extensive viscous flow. The superposed younger craters have not relaxed to the same degree. *Voyager 2 image 20638.31 PICNO 0544J2-001.*

laxed as result of slow viscous creep in the cold ice surrounding them.

The viscous degradation of craters can be analyzed by more sophisticated means than that leading to Equation 8.4.3. In particular, the effect of layers in the planet's surface can be investigated. However, all of the analytic solutions similar to Equation 8.3.4 incorporate the implicit assumption that the viscosity  $\eta$  is independent of stress  $\tau$ —that is, that planetary surface materials can be represented as Newtonian viscous substances. While this is approximately true in many situations, more recent studies have focused on numerical solutions using realistic stress-dependent viscosities. Unfortunately, each numerical solution is unique and cannot be applied to other craters. For this reason I will outline the analytic solutions below, since they do give some insight into how layering affects the relaxation of various crater sizes. The effect of a nonNewtonian rheology may be approximated by assigning a Newtonian material a small yield stress, similar to the use of a Bingham material to approximate a power-law fluid discussed in Section 8.3.5.

Most craters are axially symmetric, greatly simplifying the analysis of crater relaxation. Any axially symmetric crater profile can be represented as an integral over a series of  $J_0$  Bessel functions, just as any linear function can be represented as a Fourier integral over sine and cosine functions. The advantage of decomposing crater profiles into  $J_0$  Bessel functions is that each function relaxes self-similarly. Thus, a given  $J_0$  component remains a  $J_0$  component as it relaxes: only its amplitude decreases. This relation remains true for arbitrary layering in the subsurface. It is violated only when the viscosity depends upon stress.

A convenient analytic representation of a simple crater profile is given by

$$z(r) = \frac{HD_a}{128} \int_0^\infty (kD_a)^3 \exp(-k^2 D_a^2/16) J_0(kr) dk \quad (8.4.5)$$

where  $z(r)$  is the depth below the surface at radius  $r$  from the crater center and  $k$  is the wavenumber, which is inversely related to wavelength  $\lambda$  by  $k = 2\pi/\lambda$ . This equation represents a rimmed crater of depth  $H$  and interior diameter  $D_a$  (diameter at the preexisting surface, not rim-to-rim diameter) with an approximately parabolic profile. The volume of the rim equals the volume of the crater bowl.

Equation 8.4.5 is a special case of a general relation. Any arbitrary crater profile  $z(r)$  is related to a function  $\xi(k)$  of wavenumber  $k$  by a *Hankel transform* given by

$$z(r) = \int_0^\infty \xi(k) J_0(kr) dk \quad (8.4.6)$$

In Equation 8.4.5 the function  $\xi(k)$  is given by

$$\xi(k) = \frac{HD_a^2}{128} (kD_a)^2 \exp(-k^2 D_a^2/16) \quad (8.4.7)$$

More generally, the function  $\xi(k)$  is determined from  $z(r)$  by an integral similar to Equation 8.4.6

$$\xi(k) = \int_0^\infty z(r) J_0(kr) r dr \quad (8.4.8)$$

Since each Bessel component  $J_0(kr)$  relaxes exponentially at a rate that depends upon the wavenumber  $k$ , the crater profile as a function of time is given by another integral similar to that of Equation 8.4.6:

$$z(r, t) = \int_0^\infty \xi(k) \exp(-t/t_R) J_0(kr) k dk \quad (8.4.9)$$

where  $J_0(kr)$  in Equation 8.4.6 has merely been replaced by a dying exponential times itself,  $\exp(-t/t_R) J_0(kr)$ . In this expression the relaxation time  $t_R$  is a function of wavenumber  $k$ , so it cannot be taken out from under the integral. At time  $t = 0$  Equation 8.4.9 is equal to Equation 8.4.6 so that  $\xi(k)$  must be interpreted as the Bessel transform of the *initial* crater profile. As time goes on, the individual components of the Bessel function die away, and as  $t \rightarrow \infty$  all components vanish and  $z(r, t) \rightarrow 0$ : the crater form approaches a flat plain, as expected.

The utility of Equation 8.4.9 is that the function  $t_R(k)$  is easily determined for a variety of rheological structures in the planetary surface underlying the crater. Thus, for a crater in a surface whose viscosity is uniform down to arbitrarily large depths, the relaxation time is

$$t_R = \frac{2\eta k}{\rho g} = \frac{4\pi\eta}{\rho g \lambda} \quad (8.4.10)$$

This expression is functionally identical to the relaxation time previously derived from elementary considerations (Eq. 8.4.4). If the wavelength  $\lambda$  in Equation 8.4.10 is replaced by twice the diameter,  $2D_a$ , then the two equations become almost numerically equal.

If the rheological structure of the planet's surface is a layer of depth  $d$  overlying a *rigid* interior, as might occur for an ice crust overlying a rocky mantle, the relaxation time is

$$t_R = \frac{2\eta k}{\rho g} \left( \frac{\cosh^2(kd) + (kd)^2}{\cosh(kd) \sinh(kd) - kd} \right) \quad (8.4.11)$$

Topographic features of short wavelength (large wavenumber) in such a layered viscosity structure relax at the same rate as in a uniform viscosity mantle. Long wavelength features ( $\lambda$  comparable to or larger than  $d$ , or, equivalently,  $kd$  greater than 1) relax more slowly. The rigid bottom in this case inhibits the viscous flow over long dis-

tances and tends to preserve features with wavelengths much longer than the layer depth.

In a planet that has a viscous layer of depth  $d$  overlying an *inviscid* interior, as might occur for a stiff lithosphere overlying a much more fluid mantle, the relaxation time is

$$t_R = \frac{2\eta k}{\rho g} \left( \frac{\sinh^2(kd) - (kd)^2}{\cosh(kd) \sinh(kd) + kd} \right) \quad (8.4.12)$$

Topographic features of short wavelength in this rheologic structure again relax at the same rate as in a uniform viscosity mantle. However, in contrast to the case of a rigid bottom, long wavelength features tend to relax *faster* than in a uniform mantle.

Finally, a remarkable solution has been found (Brennen, 1974) for the case of a viscosity that either increases or decreases exponentially with depth according to the law

$$\eta(z) = \eta_0 \exp(\pm z/d) \quad (8.4.13)$$

where the + sign corresponds to a viscosity that increases with depth  $z$  and the - sign corresponds to a viscosity that decreases with increasing depth. The dimension  $d$  is the distance over which the viscosity changes by a factor  $e$ .

The importance of this solution is that the temperature increases linearly with increasing depth in most planetary bodies and, as Equation 8.4.1 shows, this leads to an exponential rate of viscosity decrease with depth for a uniform material composition. Equation 8.4.13 thus provides a reasonably good approximation to Equation 8.4.1 in the vicinity of the surface and describes the common lithospheric structure of a stiff outer layer overlying a more fluid interior.

The relaxation time for this rheologic model is a somewhat more complex function that that derived previously:

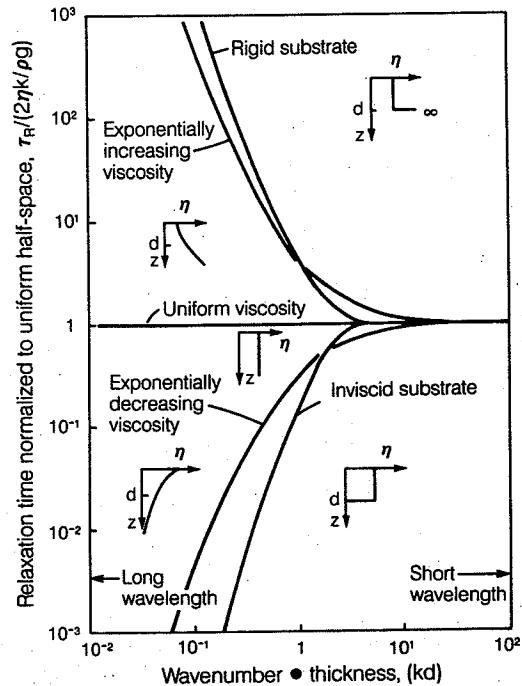
$$t_R = \frac{2\eta_0 k}{\rho g} \left( \frac{1 \pm 2\gamma}{4kd\gamma} [1 + (\gamma/kd)^2 (1 \pm 2\gamma)] \right) \quad (8.4.14)$$

where the function  $\gamma$  is given by

$$\gamma = (2)^{-3/2} [1 + 4(kd)^2 + \sqrt{16(kd)^4 + 24(kd)^2 + 1}]^{1/2} \quad (8.4.15)$$

As might be expected, the relaxation time is the same as for the uniform viscosity mantle in the short wavelength limit. In the long wavelength limit, it behaves as a viscous layer overlying a rigid base when viscosity increases exponentially with depth and as a viscous layer overlying an inviscid mantle when the viscosity decreases with depth.

Figure 8.28 shows graphically how the relaxation time depends upon wavenumber times the layer depth,  $kd$ , for all of the models discussed. It

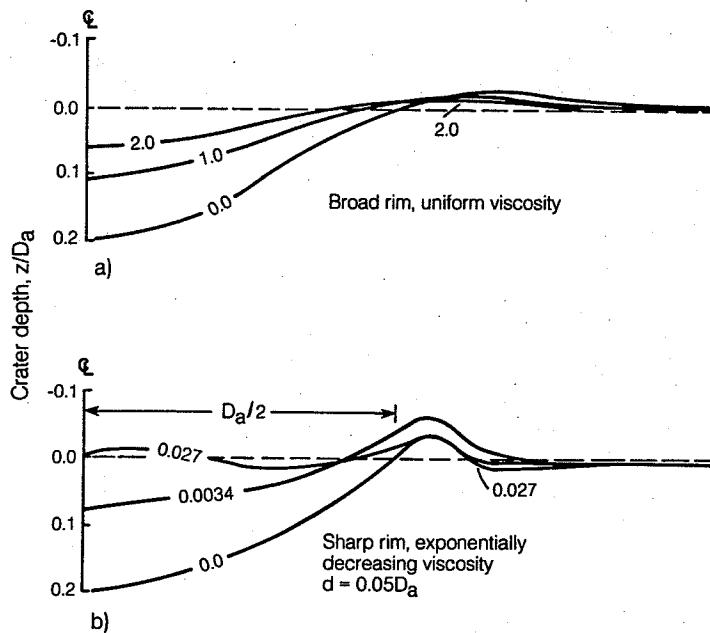


**Fig. 8.28** Relaxation time of topography as a function of its wavenumber  $k$  for several different dependences of viscosity upon depth. The relaxation time is normalized by the rate of relaxation in a uniform viscosity half space. Shown are the relaxation times for topography on a viscous layer of depth  $d$  overlying a rigid substrate, for viscosity increasing exponentially with depth over an  $e$ -folding distance  $d$ , for viscosity decreasing exponentially with depth over the same distance, and for a viscous layer of thickness  $d$  overlying an inviscid substrate. The effect of the viscosity's depth dependence is apparent only at wavenumbers such that  $kd < 1$ .

is clear that all the rheologic laws give the same result in the short wavelength limit ( $kd \rightarrow \infty$ ). The relaxation time is always shorter than the uniform viscosity case when viscosity decreases with depth, but it is always longer than the uniform viscosity case when the viscosity increases with depth.

More complex rheologic structures may be envisioned in which the surface is underlain by more than one layer of different viscosity. There are straightforward techniques for computing the relaxation time in these cases, but none is particularly simple and the pursuit of these cases is left to the interested reader. A reference to these techniques is at the end of the chapter. The cases already examined encompass most of the important qualitative behavior of viscous relaxation in a viscous substrate.

Figure 8.29a illustrates the sequence of crater relaxation in a uniform viscous medium. The initial crater profile is given by Equation 8.4.5. As



**Fig. 8.29** Profiles of craters in a viscous medium at several different times. (a) is a broad-rimmed crater (Equation 8.4.5) in a medium of uniform viscosity. The curves are labeled in units of relaxation time  $t_R$  (Equation 8.4.4). (b) is a sharp-rimmed crater in a medium with viscosity decreasing exponentially with depth. Time labels are the same as in (a). Note the more rapid relaxation and the pronounced upbowing of the crater floor. (b) is drawn after Passey (1982).

time progresses, the crater shallows markedly and its rim generally subsides, although a brief period of rim *uplift* may occur shortly after relaxation begins. Figure 8.29b illustrates the relaxation of a sharp-rimmed crater on a surface with viscosity decreasing exponentially with depth. The initial rim uplift is more striking in this case and a shallow moat eventually develops outside the crater rim. The rim structure in both cases lasts far longer than the crater bowl because of its shorter topographic wavelengths.

The crater floor in Figure 8.29a bows upward into a central dome. Although a slight upward doming is apparent even in Figure 8.29a, this tendency is greatly exaggerated when the viscosity changes rapidly with depth, and such domed crater floors are a common feature on Ganymede and Callisto. Numerical computations using stress-dependent viscosities also indicate that crater floors tend to relax into domes before they flatten out completely.

Thus far, viscous relaxation has been discussed only in a fluid of uniform density. The end result of such relaxation is a perfectly level plain, an end that seems to be an apposite description of palimpsest craters on Ganymede and Callisto. However, the moon and terrestrial planets possess low-density crusts overlying denser mantles. The excavation of a crater in the crust thins it beneath the crater. The subsequent viscous flow in either the mantle alone or the mantle and crust together results in an uplifted mantle plug beneath the thinned crust that develops on a timescale given

by Equation 8.4.4. This relatively rapid uplift halts when isostatic equilibrium is attained; that is, when the pressure at some depth in the mantle beneath the crater's center equals the pressure at the same depth far away from the crater.

The end result of isostatic equilibrium is not, therefore, a level plain, but a depression or basin on the site of the original crater. The depth,  $H_{\text{isostatic}}$ , of this basin is given by

$$H_{\text{isostatic}} = \left( \frac{\rho_m - \rho_c}{\rho_m} \right) (d_{\text{crust}} - d_{\text{crater}}) \quad (8.3.16)$$

where  $\rho_m$  and  $\rho_c$  are, respectively, the densities of the crust and mantle,  $d_{\text{crust}}$  is the crustal thickness far away from the crater, and  $d_{\text{crater}}$  is the crustal thickness beneath the crater. The height  $d_m$  of the mantle uplift beneath the crater is simply  $d_m = d_{\text{crust}} - d_{\text{crater}} - H_{\text{isostatic}}$ .

Given enough time, viscous flow will eliminate even the mantle plug beneath the crater and level the isostatic depression. However, the timescale for such flow is much longer than that given by Equation 8.4.4 because the stress differences both are smaller and extend only over short distances beneath isostatically compensated topography. Eventually, the lower crust will flow inward and allow the mantle plug to subside, but on the moon, at least, this has evidently not occurred in over  $4 \times 10^9$  years of geologic history because the large, isostatically compensated craters are still topographic lows.

The viscous relaxation computations discussed

above implicitly assume that the planetary surface material in which the relaxing crater forms may be treated as a fluid. That is, that the Maxwell time of the material is short enough that viscous flow occurs over some interval of geologic time. Although this assumption appears to be valid on such bodies as Ganymede and Callisto, it is not valid on the moon, earth, or probably the other terrestrial planets (with the possible exception of Venus, where high surface temperatures enhance the flow of even dry silicate rocks). On these planets the surface rocks are so cold that they respond elastically to stresses over the entire length of geologic time. Nevertheless, increasing temperatures in their interiors permit flow over much shorter timescales at sufficient depths. The rheologic structure of these planets is characterized by an elastic lithosphere with a thickness ranging between tens and hundreds of kilometers, depending upon planet, that overlies a fluid interior. A sufficiently large crater may be affected by relaxation in this deeper layer and flatten until isostatic equilibrium is attained.

The elastic lithosphere is deformed and generally uplifted by the flow beneath it as isostatic rebound takes place. If the uplift is large enough, the lithosphere may fracture. An investigation of the nature of these fractures (Melosh, 1976) showed that extensional tectonic features (normal faults or grabens) develop within and around the crater. They extend radially away from the center out to distances of several crater radii. A number of such features are observed around the large lunar Imbrium basin. Cauchy Rille and the Straight Wall are well-known lunar tectonic structures that strike approximately radially to the Imbrium crater and that developed some time after it formed, presumably as the result of slow isostatic rebound. Radial grabens are observed surrounding many other large impact structures on the terrestrial planets and satellites.

On a smaller scale, a class of lunar craters known as "floor-fractured craters" exhibit systems of radial and concentric cracks or ridges upon their floors. The lunar craters Humboldt and Gassendi are examples of this type. These structures have been variously attributed to localized igneous intrusions beneath the craters (Schultz, 1976) or, more recently, to viscous relaxation as a result of local thermal anomalies near the moon's surface (Hall et al. 1981). In either case, uplift of the overlying brittle layers produces a radial and concentric pattern characteristic of doming.

Theoretical computations of the stresses in the lithosphere beneath a relaxing crater give a simple prescription for estimating when failure occurs. The failure criterion for material of negligible cohesion but about 30-degree internal friction overlying a crater of diameter  $D$  and depth  $H$  (after the rapid collapse accompanying excavation)

is

$$HD/d_{el}^2 \geq 0.5 \quad (8.4.17)$$

where  $d_{el}$  is the thickness of the elastic lithosphere. This criterion is independent of viscosity because it is assumed that relaxation is complete. In actuality the failure stresses develop slowly, on a timescale comparable to the relaxation time  $t_R(k)$  with  $k \approx 4\pi/D$ . The radial fractures may thus develop considerably later than the crater itself, as observed around the Imbrian basin.

The existence of a lithosphere thus affects the slow viscous degradation of large craters. The next chapter will show that its presence may have much more profound consequences for crater collapse on a short timescale, for the lithosphere is probably responsible for the formation of the great mountainous rings that encircle the site of large impacts on the moon, Ganymede, and Callisto.

## FURTHER READING

### *General References*

The best general overview of impact crater modification is somewhat old, but contains most of the elements of the modern views on the subject:

W. L. Quaide, D. E. Gault, and R. A. Schmidt, Gravitational effects on lunar impact structures, *Ann. N.Y. Acad. Sci.* 123: 563-572 (1965).

### *Special Topics*

Much of our understanding of the formation of both simple and complex craters is derived from the detailed geologic studies of Canadian craters by M. R. Dence, R.A.F. Grieve, and their collaborators. Their current ideas are summarized in

R.A.F. Grieve, M. R. Dence, and P. B. Robertson, Cratering processes: As interpreted from the occurrence of impact melts, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 791-814 (1977).

Morphometry, the quantitative study of impact crater morphology, was begun by R. B. Baldwin in his classic book, *The Measure of the Moon*, University of Chicago, 1963. The subject has since been vigorously pursued by a number of planetary geologists. One of the best modern summaries of this work is by a former student of Baldwin's:

R. J. Pike, Size-dependence in the shape of fresh impact craters on the moon, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 489-509 (1977).

W. Hale spent several years studying the morphometry of central peaks in craters on all the terrestrial craters. A useful sample of this work is

W. Hale and J. W. Head, Central peaks in Mercurian craters: Comparisons to the moon, *Proc. Lunar Planet. Sci. Conf. 11th*, pp. 2191-2205 (1980).

These morphometric studies yielded insight into the central-peak crater to peak-ring crater transition in

W. Hale and R. A. F. Grieve, Volumetric analysis of complex lunar craters: Implications for basin ring formation, *J. Geophys. Res.* 87, Suppl.: A65-A76 (1982).

The universality of peak-ring craters is well brought out in

C. A. Wood and J. W. Head, Comparison of impact basins on Mercury, Mars, and the Moon, *Proc. Lunar Sci. Conf. 7th*, pp. 3629-3651 (1976).

Terraces in complex crater have just recently received special attention by

S. J. Pearce and H. J. Melosh, Terrace width variations in complex lunar craters, *Geophys. Res. Lett.* 13: 1419-1422 (1986).

Pit craters were first discovered in abundance on Ganymede and Callisto. The most complete description of their form is found in

Q. R. Passey and E. M. Shoemaker, Craters and basins on Ganymede and Callisto: Morphological indicators of crustal evolution, in D. Morrison (ed.), *Satellites of Jupiter*, University of Arizona Press, Tucson, pp. 379-434 (1982).

The geology of terrestrial impact craters and the mechanics of their formation is the subject of a recent review by an outstanding investigator of impact craters. The review contains further references to the original literature.

R.A.F. Grieve, Terrestrial impact structures, *Ann. Rev. Earth and Planet. Sci.* 15: 245-270 (1987).

The phenomenology of impact crater collapse is discussed in two papers. The first deals solely with wall slumping and floor uplift in a plastic substrate:

H. J. Melosh, Crater modification by gravity: A mechanical analysis of slumping, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.),

*Impact and Explosion Cratering*, Pergamon Press, New York, pp. 1245-1250 (1977).

The second paper introduces viscosity to plasticity and describes an elementary Bingham plastic model of collapse:

H. J. Melosh, A schematic model of crater modification by gravity, *J. Geophys. Res.* 87: 371-380 (1982).

The "frozen tsunami" theory of impact crater collapse, a model that proposes that lunar crustal rocks are somehow fluidized following a large impact, is discussed in detail in the next chapter. However, since the theory is supposed to apply to the peak-ring craters discussed in the present chapter, I refer the reader to the paper here:

R. B. Baldwin, the tsunami model of the origin of ring structures concentric with large lunar craters, *Phys. Earth and Planet. Interiors*, 6: 327-339 (1972).

My theory of acoustic fluidization is an attempt to establish a physical foundation for the successful Bingham phenomenological model of crater collapse. The most readable exposition is probably in

H. J. Melosh, Acoustic Fluidization, *American Scientist* 71: 158-165 (1983).

Study of the slow viscous relaxation of craters and other landforms has a long history that has recently been revitalized by the discovery of relaxed craters on the icy satellites of planets in the outer solar system. Some of the earliest work by Z. F. Daneš was unfortunately published in a 1962 research report that is now nearly unobtainable. Independent work somewhat later combined theoretical and experimental investigation:

R. F. Scott, Viscous flow of craters, *Icarus* 7: 139-148 (1967).

More recent work permits study of craters with arbitrary initial profiles and incorporates the effect of crustal layers with a different density than the underlying mantle. A good example of this work is

S. C. Solomon, R. P. Comer, and J. W. Head, The evolution of impact basins: Viscous relaxation of topographic relief, *J. Geophys. Res.* 87, B5: 3975-3992 (1982).

## IX. MULTIRING BASINS

### 9.1 DISCOVERY AND TYPE EXAMPLES ON THE MOON AND CALLISTO

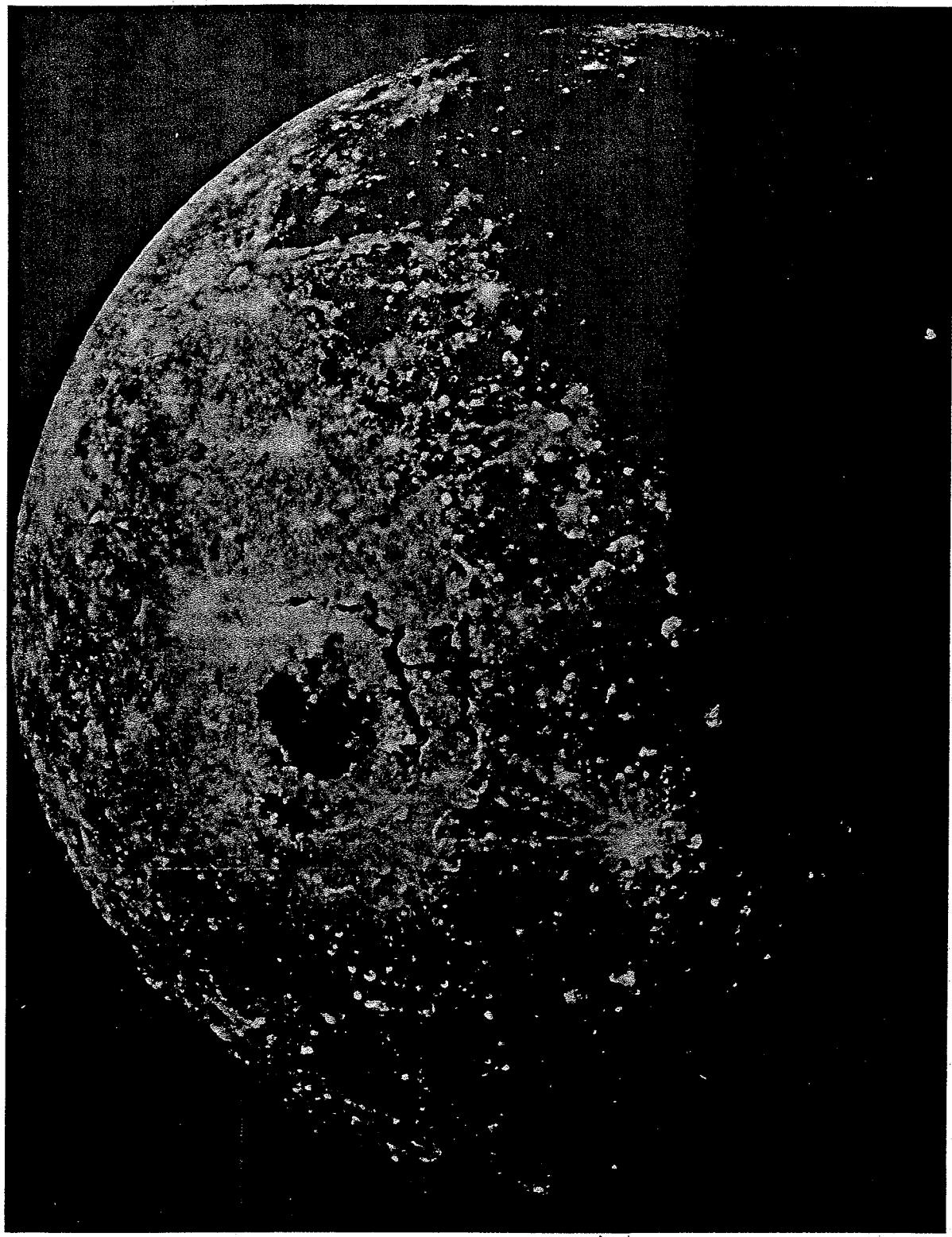
The discovery of multiring basins on the moon was very much a case of "not seeing the forest for the trees." Not recognized until as recently as 1962, multiring basins are the largest crater-related features on the moon and appear to dominate its tectonic and stratigraphic framework. Earlier workers (back to G. K. Gilbert in 1893) had described the system of ridges and troughs radial to the Imbrium basin, but none had recognized the significance of the prominent circular rings that surround Imbrium and other basins until 1962 when W. K. Hartmann and G. P. Kuiper discovered the Orientale basin from a study of rectified photographs of the Moon's western limb. Earlier workers missed the association between rings, interior plains (mare on the moon's nearside), and radial striations largely because they had concentrated on refining the moon's topography in ever greater detail—the largest-scale associations escaped their notice completely!

Hartmann and Kuiper (1962) originally defined multiring basins (or simply "basins" as they were then designated, to separate them from craters) as "large circular structures with not just one rim but an additional raised ring or rings and a system of radial furrows." This definition is now in need of modification to separate multiring basins from peak-ring craters, which fulfill the original definition, but which now appear to be distinct from the multiring structures that Hartmann and Kuiper first recognized. Peak-ring craters, discussed in the previous chapter, are an elaboration of central peak craters, whereas the great ring scarps observed on the moon form by a very different process. A better distinction, discussed below, requires that a multiring basin possess at least two asymmetric, scarped rings, one of which may be the original crater rim.

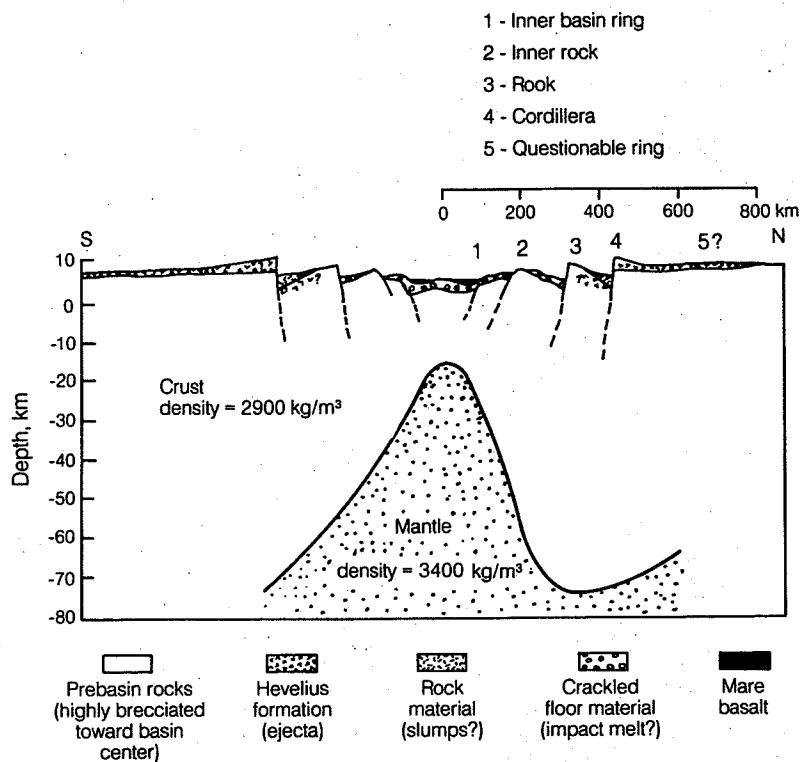
Orientale is the youngest and best preserved of the lunar basins and has therefore had a major influence on our interpretation of basins. Illustrated in Figure 2-8<sup>7</sup> and, in full sunlight but at poorer resolution, in Figure 9.1, Orientale looks like a giant bullseye on the moon's western limb. The basin center, which is occupied by a small, dark, smooth mare, is surrounded by at least four rings named the Inner Basin Ring (320 km in diameter), the Inner Rook Mountains (480 km in diameter), the Rook Mountains (620 km in diameter), and the Cordillera Mountains (920 km in diameter). Dark patches of mare basalt were locally extruded at the foot of the Rook and Cordillera escarpments. A poorly defined outer ring, 1300 km in diameter, has been suggested by some observers. The escarpments of the Rook and Cordillera Mountains are the most prominent, standing up to 6 km high. These escarpments face inward toward the basin center. The backslopes of these escarpments are gentle, producing an overall stepped profile to the basin (Fig. 9.2). The rings are not exactly circular: they are neither exactly concentric nor do they form exact arcs, so it is difficult to define a precise diameter, which accounts for variations of 10 to 20 km between different diameter estimates on even the best defined rings. The crests of the escarpments are irregular in detail, showing frequent embayments or salients. Their form can best be described as polygonal. Furthermore, the well-developed ring structures illustrated on the eastern side of Orientale seem to break down on the poorly imaged western side. On this side the rings degenerate into a zone of disorganized blocks among which rings are difficult to recognize.

The grooved and lineated ejecta blanket surrounding all lunar basins, the associated secondary craters, and samples returned from the moon's surface have convinced all but the most stubborn planetary geologists that lunar basins are of impact origin. However, if multiring basins evolved from craters, where is the original rim of the crater? Orientale, the best-exposed lunar basin, has had *each* of its rings cited as the original crater rim by one geologist or another. Current opinion leans toward associating the Rook Mountains with the original rim. The interior rings would then be analogous to peak rings, perhaps resulting from oscillation of fluidized rock debris in the crater's interior, whereas the outer Cordillera escarpment is the result of some new process associated with very large craters.

My own opinion is based on the photograph shown in Figure 9.1, which has not circulated widely among Western planetary geologists because it was made by the Soviet Zond 8 spacecraft. This figure shows a small dark ring of mare material in the southern part of Orientale that is intersected by the Rook Mountain ring. It is evidently the trace of a small basin that predated Orientale and whose presence guided the ascent of mare lavas extruded after the formation of Ori-



**Fig. 9.1** The Orientale multiring basin on the moon in full sunlight. Narrow lanes of dark mare basalt lie at the bases of the Cordillera and Outer Rook ring scarps. A small dark circle of mare basalt straddles the Outer Rook ring to the southwest of the basin center. The rings are apparently disorganized to the west. USSR Zond 8 photograph, October, 1970.



**Fig. 9.2** Geologic and geophysical structure of the Orientale basin. A dense mantle plug underlies the center of the basin. The crustal thinning above this plug is evidently due to about 40 km of crustal material ejected from the crater that formed the basin. The great ring scarps shown in cross section formed during collapse of the crater. Note the 10 times vertical exaggeration necessary to show the ring scarps. *Geology after Howard et al. (1974), crustal structure after Bratt et al. (1985).*

entale. If the Rook Mountains represent the original crater rim, I would have expected this small preexisting basin to have been completely obliterated within the crater cavity, since surface material inside the crater rim is ejected. I thus believe that the original crater rim of Orientale was well within the Rook Mountain ring—perhaps it is associated with one of the inner rings. In any case, the original transient crater was probably less than 400 km in diameter.

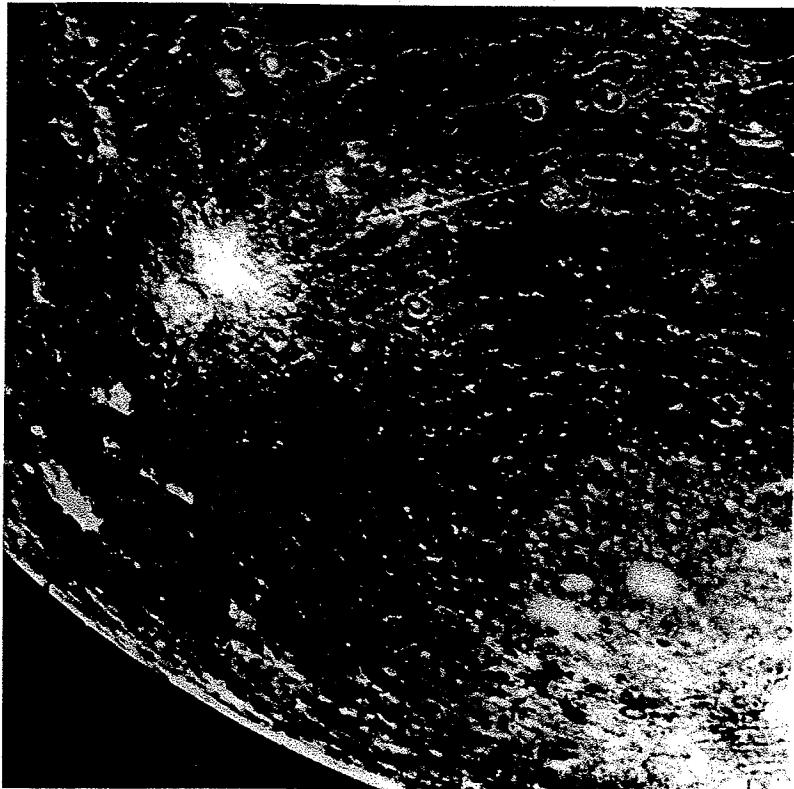
The location of the original crater rim in other lunar basins is even more obscure. There is no example in which a terraced escarpment provides a more-or-less irrefutable guide to the rim location, as we have for Compton (Fig. 8.8), Schrödinger (Fig. 8.9), or other peak ring craters. The reason for this is not clear. Perhaps the collapse of the very large craters that formed the multiring basins was so violent that no recognizable rim could survive, or perhaps some aspect of the ring-forming process either aided its obliteration or preempted the formation of a terraced rim.

Furthermore, most lunar nearside basins are flooded with lava, which produces the smooth dark mare, and details of the interior rings are smothered by the lava. In many cases only the most prominent ring, usually the outer one, is clearly visible. The interior rings may be discern-

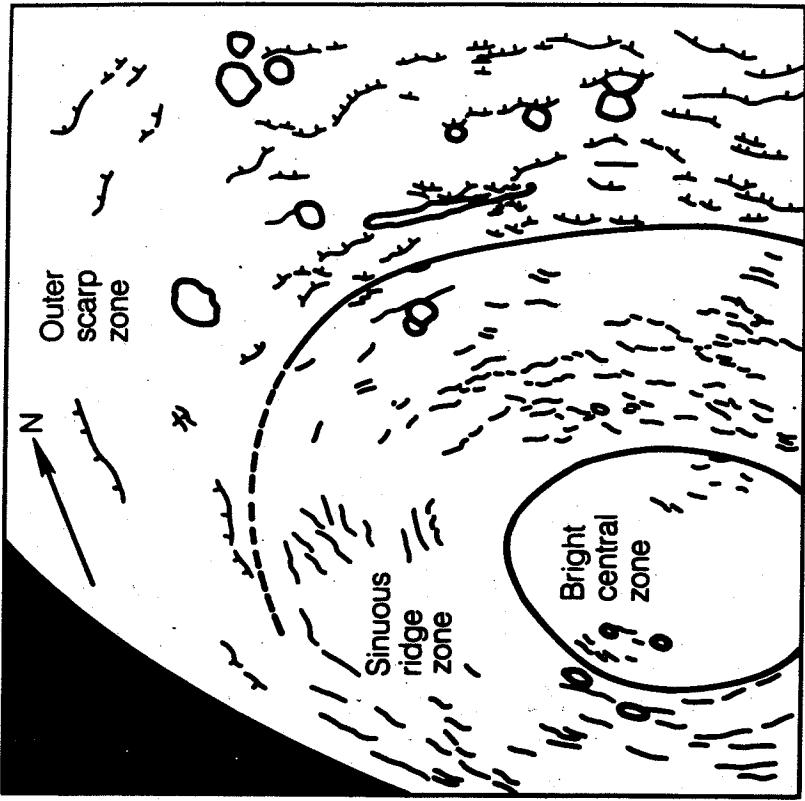
ible only as isolated hills rising above the lava or as rings of mare ridges that probably reflect the underlying structure. Lunar farside basins generally lack a lava infilling, but they have all been heavily battered by later impacts, again obliterating details of their structure.

One thing that recent research has made clear, however, is that multiring basins are distinct from the peak-ring craters discussed in the last chapter. Some earlier work tended to consider the two types gradational, and it is still possible to find lists of “multiring basins” that are primarily lists of peak-ring craters. The most obvious distinction between the two is that in peak-ring craters one or more rings form *inside* the original crater rim, whereas in multiring basins the ring scarps appear to form *outside* the original crater rim. The principal morphological difference between these ring types is that peak rings tend to have, on average, a symmetric profile, whereas Orientale’s ring profiles are distinctly *asymmetric*: they have steep inward-facing escarpments and gentle backslopes. This asymmetry is shared by all of the other prominent multiringed structures on the moon. The rims of complex craters are also asymmetric in this fashion, but they are easy to separate from rings in fresh craters because complex crater rims are terraced. Even degraded peak-ring craters and

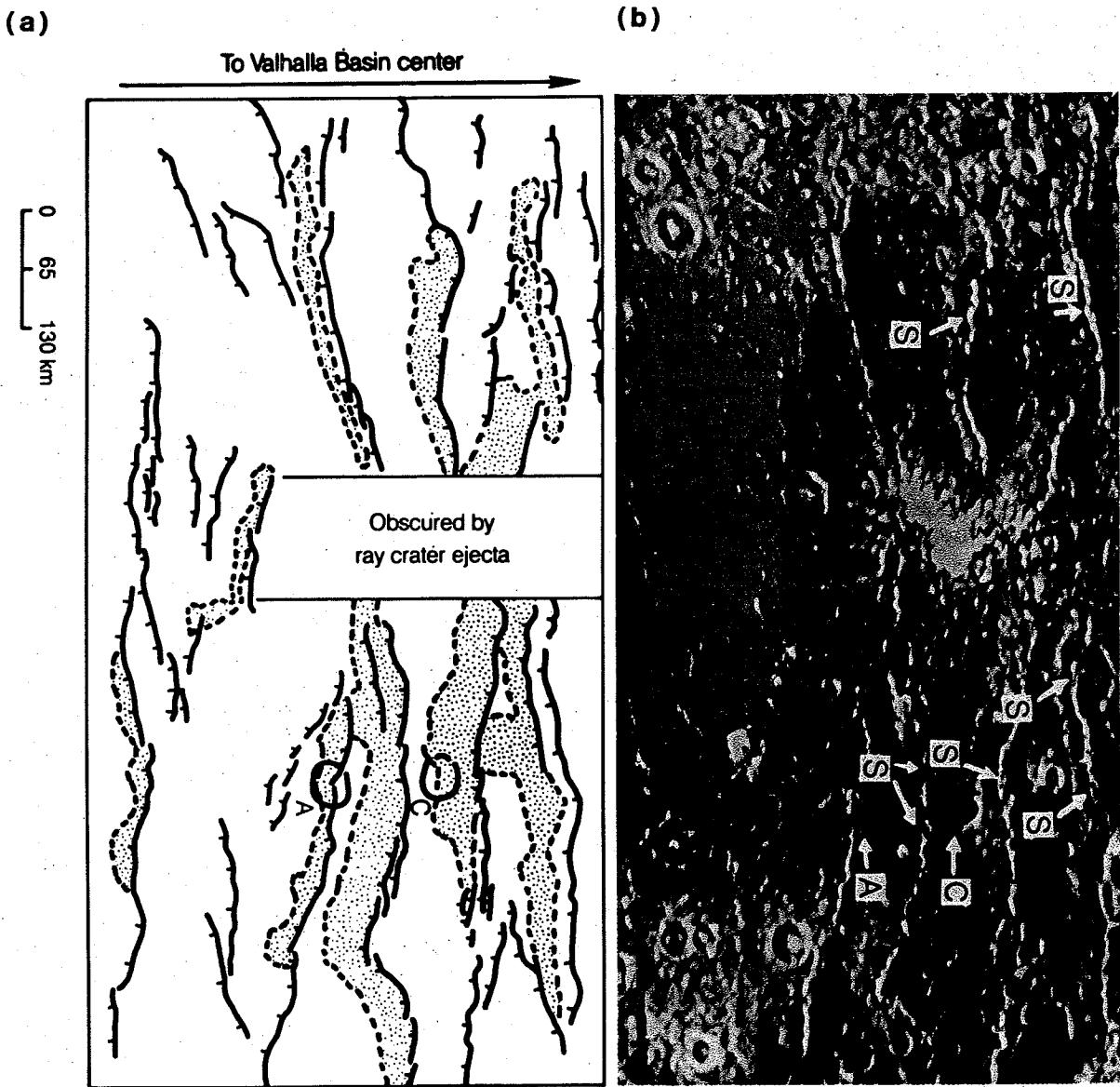
(a)



(b)



**Fig. 9.3** The Valhalla basin on Callisto. The drawing illustrates the different facies described in the text. Hatchures are on the downthrown side of scarps. A peculiarity of Valhalla is that the bright, north-facing scarps are actually facing away from the sun: their brightness is apparently due to frost on their surfaces. North-facing crater walls are also bright. The scale is indicated by the limb curvature of this 2400-km radius satellite. *After Melosh (1982b).*



**Fig. 9.4** Detail of the outward-facing scarps surrounding Valhalla Basin. The stippled areas in the sketch map are the bright terrain units that commonly lie at the feet of the escarpments. They have the same superposed crater density as Valhalla's bright center and presumably are frozen water magma extruded after the scarps formed. A and C are craters cut by the scarps. The sinuous, en echelon character of the scarps is typical of normal fault scarps in extensional terranes. *After Remsberg (1982).*

multiring basins can be separated by specifying that multiring basins must have at least *two* asymmetric escarpments, one of which may be the original crater rim.

The subsurface structure of lunar multiring basins has been probed by geophysical methods. Both gravity and topography data is available over most of the nearside basins, permitting extensive geophysical modeling of the density structure beneath the basins. Recent studies show that the center of each basin is underlain by a plug of dense lunar mantle material that intrudes through the less dense lunar crustal rocks that surround the basin and underlie the ring structures (see Fig.

9.2). The mantle plug was presumably emplaced as the transient crater, which may have penetrated most of the crust, collapsed. Since much of the overlying crustal material was ejected, the cavity could only be filled by denser mantle rocks, producing the plug and, incidentally, making the basin center a topographic low after isostatic equilibrium was attained. It is interesting to note that the diameter of the mantle plug beneath Orientale is about 600 km, smaller than the diameter of the Rook ring. This suggests that the original rim of the crater was no bigger than the Rook ring. The transient crater could have been smaller if inward transport of crustal rocks during collapse broad-

ened the mantle plug by thinning the adjacent crust.

The study of multiring basins was further advanced by the discovery of a lunar-type multiring basin on Ganymede during the 1979–1980 Voyager flybys, and by the recognition of a variant of the lunar multiring basin type on Ganymede and Callisto. Typified by the 4000-km diameter Valhalla structure on Callisto (Fig. 9.3), these multiring basins may have literally *hundreds* of rings surrounding their original impact crater. Although these rings are generally composed of asymmetric escarpments, they appear to face *outward*, not inward as lunar rings do. These outward-facing escarpments are typically a few kilometers high. They are seen to cut and displace preexisting craters (Craters A and C in Figure 9.4). A bright, apparently extrusive, deposit that obliterates preexisting topography frequently occurs at the foot of the escarpments. The crater density on this deposit is indistinguishable from that of the plains that cover the interior of the crater in the center of Valhalla, indicating that it is about the same age.

The rings of Valhalla show distinct facies, depending upon radius. The center of the structure is occupied by a bright, roughly circular patch 600 to 800 km in diameter. This bright patch may represent a crater palimpsest and should thus include both the trace of the original crater and its high-albedo ejecta blanket. No original crater rim is discernible, although a smooth plain roughly 350 km in diameter occupies the center. The bright central area is surrounded by darker terrain that is traversed by bright, roughly concentric sinuous ridges spaced 20 to 30 km apart. A few of these ridges also occur in the outer 100 km of the bright central zone. This zone of sinuous ridges (Fig. 9.3) is about 300-km wide. After a narrow transitional region in which a few grabens can be recognized, it gives way to an outer zone of outward-facing scarps. This zone is roughly 600-km wide, al-

though individual scarps may be seen as far as 2000 km from the structure's center. The outward-facing scarps are spaced at distances of 50 to 100 km.

The Valhalla structure is somewhat asymmetric in plan view: it appears to extend farther to the northeast than to the southeast. This asymmetry may reflect differences in the underlying geologic structure, as Valhalla occupies a large part of Callisto's total surface (its 4000-km diameter is comparable to the 4840-km diameter of Callisto itself). The nature of the sinuous ridges is unknown. Lighting conditions were poor for resolving them in both Voyager flybys. However, sinuous ridges apparently similar to Valhalla's were observed under low-angle lighting around the smaller Asgard basin on Callisto. The ridges appear to be flat topped, and in places a central groove or furrow is resolved within the ridge. Perhaps they are grabens or some other type of extensional structure, more or less modified by extrusion of "magma" (probably liquid water) from Callisto's interior.

Altogether, eight multiring structures have been discovered so far on Callisto. All of them are smaller than Valhalla and have fewer rings (Table 9.1). At least five poorly preserved Valhalla-type multiring basins have been identified on Ganymede (Table 9.2), which probably include the largest coherent tectonic feature on the satellite itself, the system of furrows crossing Galileo and Marius Regios. Although a number of investigators feel that these furrows are of purely tectonic origin because older maps showed them to be noncircular, a recent study using rectified maps shows that the furrows form a system that is as circular as the ring system surrounding Valhalla. The crater central to this structure has not been preserved, but younger grooved terrain occupies the site where it would normally be found. Smooth terrain within southern Galileo Regio and part of Central Marius Regio could be rem-

TABLE 9.1 *Multiring basins on Callisto*

Name or Location (Lat., Long.) if Unnamed	Diameter* (km)		Ring Spacing, (km)	
	Outermost Ring	Innermost Ring	Inner Zone	Outer Zone
Valhalla	4000	400	30	50–100
Asgard	1640	163	20–30	50–60
(−53, 36)	920	130	15	30
(45, 138)	500	—	—	—
Grimr	180	96	5	—
Alfr	163	80	7–9	—
Loni	123	25	10	—
(−41, 262)	71	35	—	—

All data are from Passey and Shoemaker (1982).

\*Diameters are  $\pm 10$  km.

TABLE 9.2 *Multiring basins on Ganymede*

A. Valhalla-type Basins*		
Name or Location (Lat, Long) if Unnamed	Outer Ring Diameter (km)	Approx. Ring Spacing (km)
Furrows in Marius and Galileo Regios	8300	22 in Marius 50 in Galileo
Furrows in Nicholson Regio	2800	100
Furrows in Perrine Regio (41, 144) (-27, 154)	1600 650 450	50 90 90

## B. Lunar-Type Basins†

Name	Ring Diameter (km)
Gilgamesh	175
Western Equatorial Basin	185

\*Data from Schenk and McKinnon (1987).

†Data from Passey and Shoemaker (1982).

nants of a dark Valhalla-like palimpsest that once surrounded the crater (Schenk and McKinnon, 1987).

## 9.2 BASIN RING SPACING

In their 1962 discovery paper Hartmann and Kuiper noted that the diameters of adjacent rings are frequently in the proportion of approximately  $1:\sqrt{2}$ . Baldwin (1963) independently remarked on a common 1:2 proportion of the same ratio. The meaning and reality of this ratio has been hotly debated by subsequent authors. Opinion ranges from the position that the ratio of adjacent ring diameters is " $\sqrt{2}$  plus or minus nothing" to doubts that the  $\sqrt{2}$  ratio has any significance at all.

One factor that has clouded this issue in the past is the inclusion of a large number of peak-ring craters in the sample of multiring basins. It is well established that the ratio of the peak-ring diameter to the crater diameter is roughly 0.5 on the earth, the moon, Mercury, and Mars (Sec. 8.3.1 and Fig. 8.9). The diameters of basin rings are also difficult to define precisely, introducing the possibility of an unconscious bias favoring the selection of ring diameters whose ratio comes close to the magic  $\sqrt{2}$  number. Some studies have even used the  $\sqrt{2}$  ratio as a guide to search for otherwise "hidden" rings.

Table 9.3 lists the ring diameters of the best-es-

TABLE 9.3 *Lunar Multiring Basins*

Name	Ring Diameters*				
	(km)				
Imbrium†	670	970	<b>1340</b>		
Orientale‡	320	480	<b>620</b>	930	1300
Nectaris†	400	600	<b>840</b>		
Smythi†			<b>450</b>	810	
Crisium†			<b>450</b>	670	
Humorum†‡			<b>270</b>	<b>410</b>	560
Moscoviense†			<b>205</b>	<b>410</b>	700
Hertzsprung†			<b>285</b>	440	550
Grimaldi‡			<b>220</b>	410	

\*Diameters of the most prominent ring are in boldface.

†Wood and Head (1976)

‡Hartmann and Wood (1971)

tablished lunar multiring basins. Many sources were consulted and the table includes only those rings that a majority of investigators are able to recognize. The ring diameters listed are also supported by a majority of the sources. Although there are obvious dangers in the democratic establishment of scientific truth, this is probably the best and most unbiased assessment of ring diameters presently available. Table 9.3 specifically excludes peak-ring craters, although some of the smaller basins are considered borderline by a few planetary geologists. The frequency of occurrence of different ring diameter ratios is plotted in Figure 9.5. In this figure, the ratio is formed between the diameter of a given ring and the basin's most prominent ring (indicated in Table 9.2 by boldface type).

The statistics in Figure 9.5 are poor, but a case can be made for preferred ratios of  $1/2$ ,  $1/\sqrt{2}$ , and  $\sqrt{2}$ . This data also shows that the diameter of rings in lunar-type multiring basins increases outward in rough geometric proportion, independent of the basin diameter. This fact must be explained (or discounted) by any theory of basin ring origin. The rings of Valhalla-type multiring basins are so numerous that no regularity has yet been suggested for their diameter ratios.

## 9.3 OCCURRENCE OF MULTIRING BASINS

Multiring basins are common on the moon and Callisto, and a few have been recognized on Ganymede. Surprisingly, only a weak case can be made for their existence on the earth, Mercury, or Mars (Venus, of course, is still largely unstudied). None have been seen on the satellites of Saturn or Uranus. Many planetary scientists may take issue with this bald statement because the literature is full of reports of multiring structures on Mercury and Mars. However, these reports are usually based on a confusion between peak ring craters,

## IMPACT CRATERING

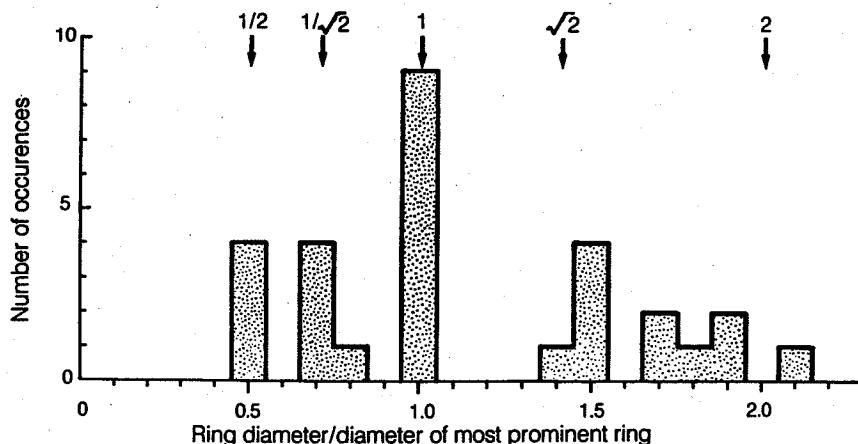


Fig. 9.5 Histogram of ring diameter ratios for well-measured rings surrounding lunar multiring basins. Preferred basin ring spacings are  $1/2$ ,  $1/\sqrt{2}$ , and  $\sqrt{2}$ , with a weaker cluster about 2. This figure illustrates the oft-stated contention that basin-ring diameter ratios are multiples of  $\sqrt{2}$ . Data from Table 9.3.

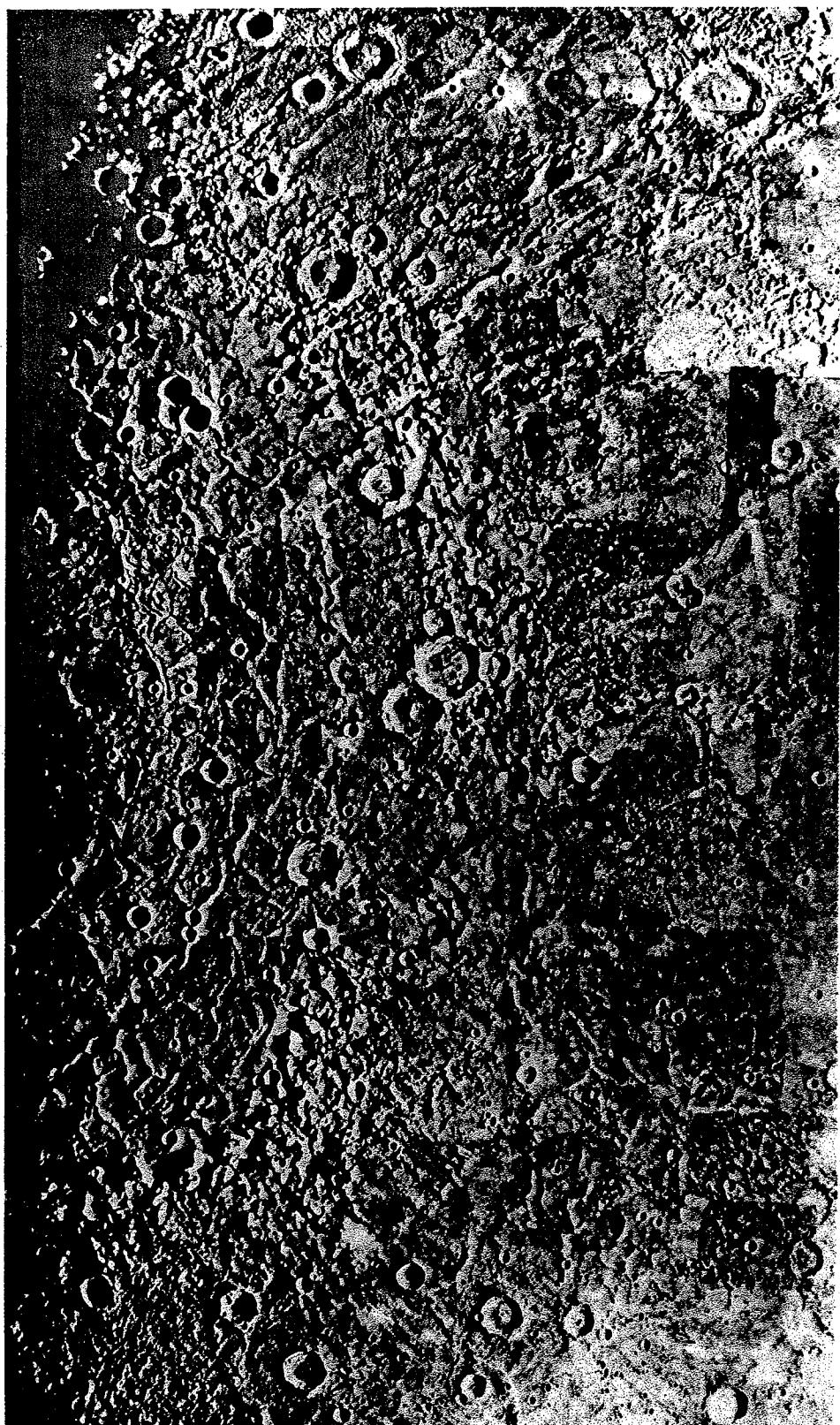
which are abundant on both planets, and the asymmetric scarp type of multiring basin. There have been many vague and poorly defined "rings" supposedly recognized around large degraded craters. Some of these "rings" are defined only by one or two topographic highs or other unusual features that seem otherwise inexplicable to the mapper and that are then raised to the status of a ring. Although some of these elusive rings may turn out to be bona fide, most are highly questionable.

One reason for the proliferation of multiple ring identifications throughout the inner solar system is that many planetary scientists have assumed that "multiring basin" is the next step in the crater size-morphology progression beyond "peak-ring crater." Along this line of thinking, craters as large as Caloris (1300 km in diameter) on Mercury or Hellas (2000 km diameter) on Mars simply *must* be multiring basins because they are so big. The direct evidence for more than one asymmetric ring around these large craters, however, is weak. This important fact was first pointed out by C. A. Wood and J. W. Head (1976), who concluded that multiring basins occur for diameters greater than 350 km on the moon, 1200-km diameter for Mars, and none at all were recognized on Mercury. Note that the moon-Mars trend of onset diameters is exactly the *inverse* of the  $1/g$  progression of the simple-to-complex or central peak-to-peak ring transitions observed in collapsed craters (Sec. 8.3 and Sec. 8.3.1).

Although no terrestrial impact structures have been shown unequivocally to be multiring basins, the Manicouagan structure in central Quebec, Canada, may be one (Grieve and Head, 1983). The inner zone of this structure is occupied by a thick sheet of shocked and melted rocks roughly 55 km in diameter. This is enclosed by a ring of

fractured rocks about 100 km in diameter, which is thought to represent the extent of the original crater. Surrounding this zone is an outer circumferential depression about 150 km in diameter, which may represent an outer ring. The ratio between the ring diameter and the presumed crater diameter is 1.5, close to the famous  $\sqrt{2}$  ratio. No outer rings have been reported surrounding larger structures such as Vredefort (160 km in diameter) or Sudbury (140 km in diameter), although both structures are highly degraded and evidence for outer rings may be difficult to find.

The largest known impact structure on Mercury, the Caloris basin, is 1300 km in diameter (Fig. 9.6). Only one prominent asymmetric scarp is recognized but its interior is flooded with smooth plains that are presumed to cover at least two interior rings. The existence of these rings is inferred from two nearly complete arcs of ridges with diameters of 800 and 1060 km on the interior plains that presumably reflect some kind of structural inheritance from underlying topographic rings. These rings are probably analogous to peak rings in large complex craters, since the interior lava deposits appear to be a few kilometers thick, at most, and steep, 6-km-high asymmetric scarps should emerge above them. One or more faint rings beyond the main scarp have been suggested at various times, but none is well defined. Another large crater, Beethoven, is surrounded by a degraded 625-km diameter scarp, but no further details are visible, so it is not clear that this structure is a true multiring basin. The smallest structure on Mercury that has been declared to be a multiring basin is the old and degraded crater Tolstoj. A 300-km diameter inner ring is supposed to define the transient crater diameter, and segments of at least three rings spaced 50 km apart are identified surrounding this inner ring (McKinnon, 1981). If these rings are real,



**Fig. 9.6** Photomosaic of the 1300-km diameter Caloris Basin, Mercury. In spite of its large size, there is little evidence for more than one asymmetric ring scarp, aside from a short, incomplete segment to the northeast. Nearly complete arcs of mare ridges in the basin interior may overlie symmetric rings of the peak-ring type.

their diameter ratios badly violate the  $\sqrt{2}$  spacing rule.

On Mars the three largest impact structures are Argyre (1200 km in diameter; see Figure 9.7), Isidis (1900 km in diameter), and Hellas (2000 km in diameter), of which Argyre is the best pre-

served. Argyre's prominent main scarp is clearly asymmetric and radially lineated by troughs and ridges. A poorly exposed interior ring some 560 km in diameter can be recognized, but its nature is unknown. A number of rings beyond the main scarp have been suggested by discontinuous con-



**Fig. 9.7** The 1200-km diameter Argyre basin in the southern highlands of Mars. Like the Caloris basin on Mercury, Argyre shows little evidence for more than one asymmetric ring scarp. *Viking photomosaic NASA S.76-27774.*

centric ridges, although some of these may be erosional features. Nevertheless, most planetary geologists feel that Argyre is probably a true multiring basin. Neither Isidis nor Hellas show better-defined multiple ring scarps, so the entire existence of multiring basins on Mars is open to question.

A lunar-type multiring basin named Gilgamesh lies within the grooved terrain on Ganymede (Fig. 9.8). A central smooth depression 150 km in diameter is believed to correspond to the central pit observed in other large craters on Ganymede. The pit is surrounded by a poorly defined 175-km-diameter ring that may be the rim of the crater that formed the basin. Several discontinuous inward-facing concentric scarps up to 1.5 km high surround the basin, defining an outer ring with a diameter of about 550 km (Passey and Shoemaker, 1982). Ring-diameter ratios are clearly far from  $\sqrt{2}$ . A second large basin, called the Western Equatorial Basin, possesses an asymmetric inward-facing ring scarp 185 km in diameter and less than 600-m high. Since there is only one scarp, it is unclear whether or not this is another multiring basin. A number of Valhalla-type basins have been recognized on Ganymede. They were discussed at the end of Section 9.1 and are listed in Table 9.2.

The Valhalla basin and similar structures on Callisto have already been discussed in Section 9.1. No lunar-type multiring basins have been recognized on this satellite.

#### 9.4 THEORIES OF MULTIRING BASIN ORIGIN

Numerous theories to explain the formation of basin rings have been proposed since multiring basins were recognized in 1962. Nearly all of these theories apply only to the lunar type of multiring basin, partly because Valhalla-type basins were not known until the 1979–1980 Voyager flybys of Jupiter and partly because many planetary scientists consider the Valhalla-type to be a distinct structure that forms by a different process than the lunar-type multiring basins. Only one theory, ring tectonics, provides a common explanation for the two basin types. Although ring tectonics has gained many adherents, it is by no means universally accepted, and much controversy still surrounds the origin of multiring basins. For this reason I will briefly review a number of other theories that have been proposed before proceeding with a more detailed exposition of ring tectonics. Interested readers will find references to further work on these other theories at the end of the chapter.

##### 9.4.1 The Volcanic Modification Hypothesis

One of the earliest suggestions by one of the co-discoverers of multiring basins (Hartmann and

Yale, 1968) was that multiring basins start like any ordinary large crater. At some later date subsurface melting or intrusion takes place beneath the crater and subsidence occurs along great ring faults surrounding it. These ring faults may be analogous to the cone sheets that form during caldera collapse of large volcanoes. The lava that wells up along these faults floods the basin floor and produces an interior mare. In this theory the inward-facing scarps that define the exterior multiple rings are the surface traces of the ring faults (Fig. 9.9).

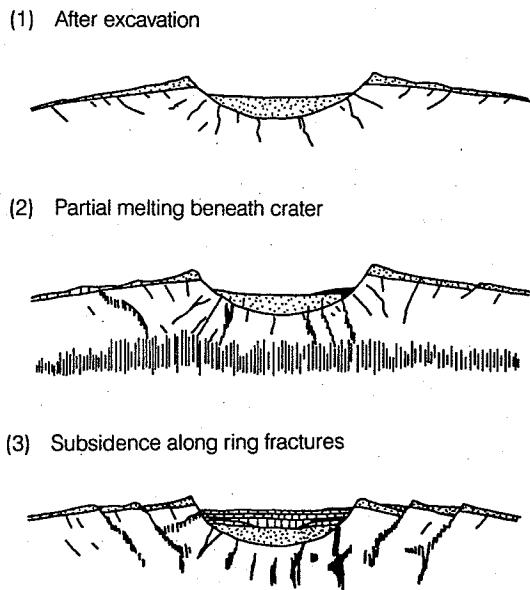
The major difficulty with this theory is that geologic evidence indicates that the multirings form very quickly after the crater is excavated. The interior of Orientale is filled with a fissured deposit that is interpreted as impact melt or Suevite. This deposit laps up against the edges of the Rook escarpment and buries associated structures, but is not itself deformed. This shows that the rings were in existence before the Suevite hardened—perhaps only minutes or hours after the basin itself was excavated. Even more telling, parts of the Cordillera escarpment are overlapped and buried by *ejecta* deposits. This observation proves that the Cordillera escarpment developed *before* the ejecta arrived—a time interval of about 13 minutes after the crater rim formed. This is simply not enough time for magmatic intrusions to be emplaced beneath the crater floor.

One way in which magma could arise nearly instantaneously beneath the crater floor is by pressure-relief melting as a result of the rapid removal of the overburden during crater excavation. However, pressure-relief melting, which is responsible for the generation of midoceanic ridge basalts on earth, is very inefficient beneath craters on the moon, partly because of the moon's low gravity and hence small pressure relief for a given uplift and partly because the uplift distance beneath even the largest lunar craters is less than about 100 km, far smaller than the distance over which terrestrial convection currents rise. Although pressure relief melting beneath large lunar craters has often been suggested as the source of mare lavas, closer examination shows it to be ineffective. The lunar mantle would have to be nearly at the melting point for pressure-relief melting to be significant, and if it is this close to the melting point, then there is no need to call upon pressure relief to generate basaltic magmas.

The volcanic modification hypothesis thus does not provide a good explanation of multiring basin formation. Historically, this theory's qualitative agreement with the facts and the observed common association of mare basalts with multiring basins led to an initially warm reception, which has subsequently cooled in the light of better data. In particular, several large farside basins seem to lack mare basalt deposits altogether. Mare basalt extrusion now seems to be mainly correlated with crustal thickness: The nearside basins are gener-



**Fig. 9.8** Gilgamesh Basin on Ganymede. The central smooth depression is 150 km in diameter. It is surrounded by a lunar-type inward-facing scarp 550 km in diameter and up to 1.5 km high. *Voyager 2 image 20638.14, PICNO 0527J2-001.*



**Fig. 9.9** The volcanic modification theory of basin-ring formation. The basin begins as a normal large crater. Melting beneath the basin at a later date generates mare basalts that flood the basin and cause subsidence along ring fractures similar to terrestrial caldera collapse. *After Hartmann and Yale (1968).*

ally flooded with basalt because the nearside crust is thin. A variation of the volcanic modification hypothesis, however, may apply to a number of smaller craters on the moon called "floor fractured craters" that appear to have been modified by magmatic intrusions beneath their floors.

This may be the place to note that a number of Soviet planetary scientists still doubt the impact origin of multiring structures and instead attribute their origin to the collapse of giant calderas. When the Valhalla structure on Callisto was first discovered, at least one serious comparison was made to the "ice calderas" of Iceland and Antarctica. Ice calderas develop on the surface of an ice sheet that covers the site of a volcanic eruption. The hot lava melts the lower part of the ice sheet away and produces a multiringed collapse structure (Wood, 1981). Such theories are presently unacceptable to the general mass of planetary scientists.

#### 9.4.2 The Megaterrace Hypothesis

J. W. Head (1974, 1977) noted that the Cordillera ring of Orientale is located about 1.5 times farther from the basin center than the Rook ring, which Head accepts as the original rim of the crater that formed Orientale. This is about the same distance from the center as the outer limit of structural uplift beneath the crater rim (see Sec. 6.2). Head proposed that the uplifted rim constitutes a zone of weakness that localizes interior collapse and

produces a very wide terrace or "megaterrace" whose formation is otherwise mechanically similar to the much smaller terraces that form around the peripheries of complex craters (see Sec. 8.3.3). The inward-facing scarp is interpreted as the outer edge of the weakened zone.

A similar suggestion was made by S. K. Croft (1981), who proposed that the strength discontinuity that localizes ring formation is not the edge of the uplifted zone, but is instead the limit to which rock surrounding the crater is brecciated by stress waves (the "strength crater"). Neither theory explains how the existence of a strength discontinuity produces the observed sharp inward-facing scarp, nor how more than a single ring can form outside the initial crater. These theories thus do not address Valhalla-type multiring basins.

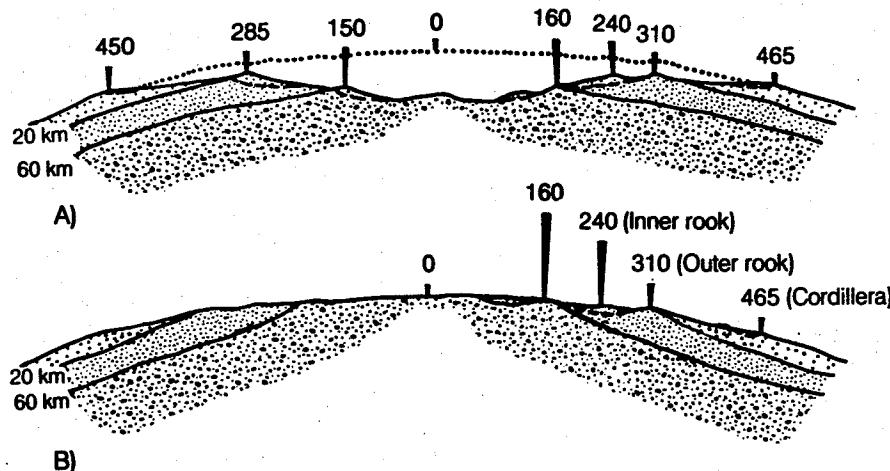
In the megaterrace hypothesis, multiring basin development is an inevitable consequence of increasing crater size. The theory cannot account for the low abundance (or, possibly, the absence!) of multiring basins on Mercury and Mars. This hypothesis was developed at a time before peak-ring craters were clearly differentiated from multiring basins. The sequence of central peak crater, peak-ring crater, and, finally, multiring basin, was then thought to be universal. The more recent findings described in the previous section make a universal sequence of this kind improbable and thus offer little support for the megaterrace theory.

#### 9.4.3 The Nested Crater Hypothesis

It has been noted (Sec. 5.6.2) that strong layers in the target produce benches or ledges in the wall of any crater that penetrates these layers. This observation prompted C. A. Hedges and D. E. Wilhelms (1978) to suggest that basin rings begin as ledges formed at the horizons of major strength increases in the lunar crust or as material slumped from one of the ledges. After excavation, isostatic uplift of the ledges or terraces results in concentric mountain rings inside the crater (Fig. 9.10). The outer basin ring is considered to be the rim of the original crater in this theory.

The major objection to this theory is that it fails to explain the observed regularities in basin-ring spacing. In the nested crater model, ring spacing is related to the thickness of layers in the lunar crust and is not proportional to the crater diameter. The nested crater model thus cannot explain the  $\sqrt{2}$ -ring-diameter ratio rule. If the 20-km- and 60-km-deep layers postulated in Figure 9.10 to explain the positions of Orientale's rings are global features of the moon's crust, then the distance between the rings of other basins should be nearly the same as Orientale's: about 150 km between major rings. Table 9.2 makes it clear that this is far from the truth. Since basins are widely scattered across the moon's surface, one would not

## IMPACT CRATERING



**Fig. 9.10** The nested crater theory of multiring basin formation. Layers with increasing strength in the moon produce benches or ledges within the transient crater. Slumping may modify these ledges. Isostatic uplift finally raises the crater floor and the ledges are tilted outward to form ring scarps. *After Hedges and Wilhelms (1978).*

necessarily expect the layer thicknesses to be exactly the same beneath all basins, but even nearby basins like Orientale and Grimaldi have vastly different ring spacings that require sudden and ad hoc changes in layer thickness in this model. The nested crater model has no simple explanation for the scarcity of multiring basins on Mercury and Mars, nor does it attempt to explain Valhalla-type multiring basins.

#### 9.4.4 Ring Tectonic Theory

The final theory discussed here grew out of studies of crater collapse in layered media in which the strength decreases with increasing depth. Under these circumstances one or more ring fractures arise outside the rim of the original crater (Melosh and McKinnon, 1978). The situation in which strength decreases with increasing depth is a common one near planetary surfaces where a strong elastic layer, the lithosphere, is underlain by a fluid asthenosphere.

Before proceeding with a discussion of the theory, it is important to define exactly what is meant by the lithosphere and asthenosphere in the context of crater collapse. These terms are normally used to describe regions near a planet's surface that respond in very different ways to loads that endure over significant lengths of geologic time. The lithosphere is the outer, relatively cold, portion of the planet that either deforms elastically or fails in a brittle fashion under applied stresses. Material in the deeper, warmer, asthenosphere deforms mainly by creep and allows applied deviatoric stresses to relax over time. The usual criterion for separating these two regions is that the Maxwell time  $t_M$  (Eq. 8.4.2) in the lithosphere is larger than the age of the surface, while the Maxwell time of the asthenosphere is smaller than the

age of the surface. The lithosphere–asthenosphere boundary does not necessarily coincide with the boundary between compositional layers in the planet, such as the crust–mantle interface.

On the timescale of crater collapse (roughly  $(D/g)^{1/2}$ ; see Sec. 8.1), however, the conventionally defined asthenosphere may respond elastically. Flow is only significant in material whose Maxwell time is shorter than the collapse time. This criterion defines the “asthenosphere” in the context of crater collapse. The “lithosphere” for crater collapse is the layer in which the Maxwell time is longer than the crater collapse time. Although this specially defined lithosphere is not identical to the conventional lithosphere, the thicknesses of the two may be proportional to each other.

At first sight it may seem that the crater-collapse asthenosphere, if it exists, must be liquid. Only liquids have Maxwell times shorter than a few seconds. The internal viscosity of most convecting silicate planets is roughly  $10^{21}$  Pa-seconds, which implies a Maxwell time of about 500 years, far longer than the collapse time of any conceivable crater. However, this viscosity is observed at stress differences on the order of 1 MPa. As discussed in Section 8.4, the creep rate of hot silicates is stress dependent, leading to a viscosity that depends upon stress to the inverse second or third power. The stress differences in the vicinity of a large collapsing crater may approach 1 GPa, producing effective viscosities  $10^6$  to  $10^9$  times smaller than those inferred from convective stresses. Maxwell times from hundreds of minutes down to a few seconds may thus result and collapse will proceed as if the hot silicate interior of the planet were fluid. A similar argument was made for ice in the interior of Ganymede and Callisto by W. B. McKinnon (1981).

When a crater forms in an elastic lithosphere

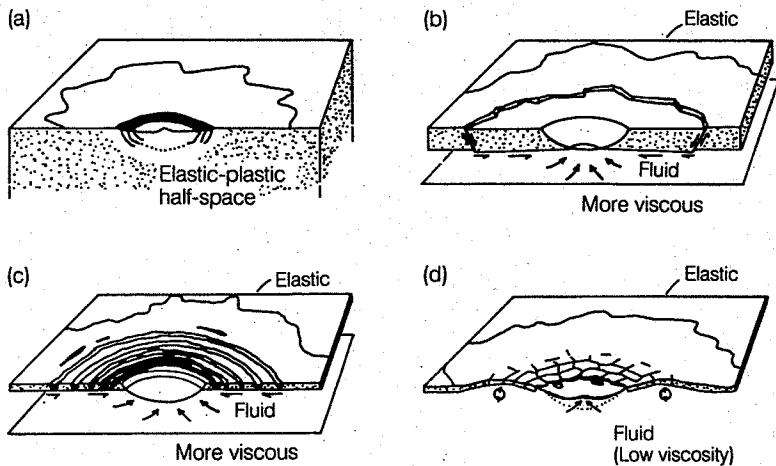
that overlies a fluid asthenosphere, one of several things may occur (Fig. 9.11). If the lithosphere is much thicker than the crater's depth, nothing special happens (Fig. 9.11a). The result is a normal crater that is either simple or complex, depending upon its size. Acoustic fluidization (Sec. 8.3.5) may facilitate the crater's collapse and produce central peaks or internal rings with symmetrical profiles, but the acoustic field's range is too short to extend into the region where multiple rings form. If the transient crater's depth is larger than the lithosphere's thickness, however, at least one ring fracture may form outside the crater (Fig. 9.11b). This fracture develops because the asthenosphere flows inward beneath the lithosphere as it intrudes and partially fills the crater. This flow exerts an inward drag on the lithosphere, which, if it is strong enough, creates an extension fracture in the lithosphere. Although this fracture may be irregular in detail, following preexisting lines of weakness in the lithosphere, it is broadly circular and resembles the asymmetric scarps around lunar craters and around Gilgamesh on Ganymede.

Recent modeling of the asthenospheric flow beneath a crater shows that an additional condition must be satisfied before multiple rings can form: the low-viscosity asthenosphere must be underlain by a higher-viscosity layer, forming a low-viscosity channel. If the lithosphere is underlain by a uniform-viscosity half space, flow beneath the crater merely tends to uplift the lithosphere and no circumferential extensional stresses develop. It is interesting to note that many rheologic models of Ganymede and Callisto call for an especially fluid layer of ice II just below the surface. This

fluid layer is underlain by stiffer phases of ice, giving rise to the type of low-viscosity channel needed for multiple-ring formation. In the case of the moon, residual heat from the early magma ocean may have created the necessary low-viscosity channel at the time that the large lunar basins formed. The rather special rheologic conditions required for multiple-ring formation may explain their apparent rarity, and certainly argue against the idea that multiring basins are an inevitable stage in the crater size-morphology progression.

A remarkable small-scale experimental illustration of the formation of ring scarps is shown in Figure 9.12. The ring is defined by a dark line surrounding the central water-filled crater of the explosion-produced Snowball Crater. Formed in water-saturated clays and sands near Alberta, Canada, this 108.5-m (rim-to-rim)-diameter crater was produced by the detonation of a 500-ton TNT hemisphere resting on the surface. The ring, which actually consists of open fissures up to a half a meter wide, developed over a period of hours after the explosion test. It appears to be the result of the inward flow of a buried stratum of water-saturated sand that was fluidized by the explosion. The presence of fluidized sand is attested to by the numerous sand blows that form bright patches on the surface outside the crater. The ring diameter in this crater is approximately twice the crater's diameter. If this crater were larger, the open fissures would have been replaced by normal faults similar to those surrounding multiring basins, since actual tension is difficult to attain on a geologic scale because of the relatively great importance of the compressive overburden stress.

According to ring tectonics, the scarp forms at



**Fig. 9.11** The ring tectonic theory of multiring basin formation. Drawings (a) through (c) illustrate the effect of decreasing lithosphere thickness, from a very thick lithosphere (a) that prevents rings from forming, to a moderately thick lithosphere (b) that produces one or a few lunar-type rings, to a very thin lithosphere (c) in which Valhalla-type ring systems develop. Recent work indicates that rings form only when the viscosity beneath the lithosphere increases again at greater depths form a low-viscosity channel. Drawing (d) illustrates the hypothetical case of a very fluid mantle underlying a thin lithosphere in which a radial-plus-concentric fracture pattern develops. After McKinnon and Melosh (1980).

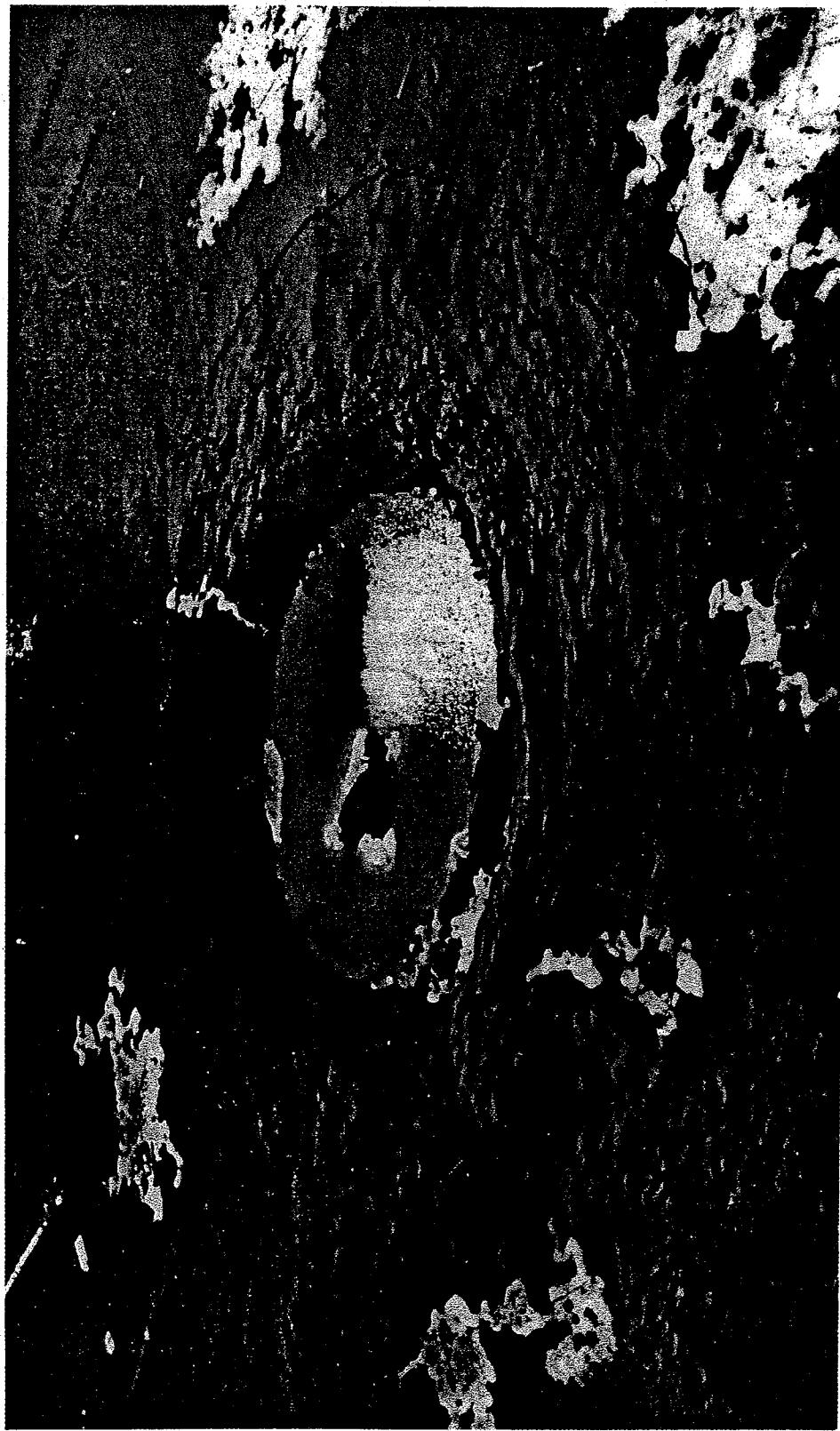


Fig. 9.12 A ring crack formed about the 108.5-m diameter Snowball explosion crater, produced by a 500-ton TNT hemisphere detonated at the ground surface. The ring crack developed after a buried, water-saturated sand layer fluidized and allowed the clay layers near the surface to slide toward the crater cavity. Photo courtesy D. J. Roddy.

a distance from the crater's rim that is proportional to the rim diameter, not the lithosphere's thickness. The scarp cannot form very close to the crater rim because shear stresses at the base of the lithosphere need to act over a broad area to build up large extensional forces within it. The scarp cannot be very far from the crater rim because a large-diameter ring fracture has so much area that it can resist the inward-directed forces. A balance is therefore struck, and the fracture, if it forms at all, occurs between about 1.3 and 1.6 crater radii from the center, depending upon details of the model. This ratio is sufficiently close to the famous  $\sqrt{2}$  ratio for agreement to be claimed. After the formation of one ring scarp, which relieves the stresses acting on the annulus of lithosphere between the crater rim and the scarp, this process may repeat itself to form a second or even third ring scarp surrounding the crater. Each new ring scarp becomes a surface of stress relief and so the diameter of the next succeeding ring is a constant factor times the diameter of the last one, as observed. The scarps near the crater face inward because of the direction of the shear stress on their bases: the inward-directed asthenosphere flow produces a couple on the lithospheric blocks broken off at a ring fracture. This couple tends to tilt the lithospheric blocks outward, resulting in an inward-facing scarp with a gentle outward-dipping backslope.

As already mentioned, the lithosphere thickness plays little or no role in determining the ring spacing. However, the thickness is the dominant factor in determining whether or not a ring forms at all. Aside from the obvious requirement that the crater must penetrate the lithosphere, the lithosphere's thickness and strength must not be too large. A simple argument involving the balance of forces against resistance will clarify this point. Suppose that a constant inward-directed shear stress  $S$  acts on the base of the lithosphere between the crater rim at radius  $R$  and a distance  $2R$  from its center. Numerical computations show that this shear stress is on the order of  $0.05 \rho g H_i$ , where  $H_i$  is the transient crater depth. The net force  $F_i$  dragging the lithosphere inward is

$$F_i \simeq 3\pi R^2 S \quad (8.4.1)$$

The resisting radial stress  $\sigma_r$  acts within the lithosphere over a cylinder of height  $h$  (roughly equal to the lithosphere's thickness) that surrounds the crater at radius  $2R$ . The net outward-directed force  $F$  is

$$F_r \simeq 4\pi Rh\sigma_r \quad (8.4.2)$$

If failure occurs when  $\sigma_r$  reaches a yield stress  $Y$ , then Equations 8.4.1 and 8.4.2 can be equated to derive a yield condition. The lithosphere thus begins to fracture when

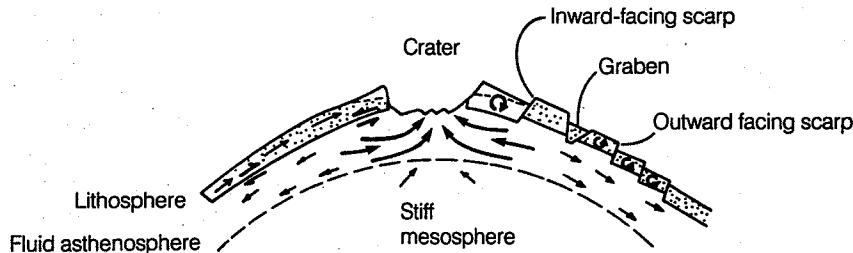
$$\frac{3 SR}{4 Y h} \gtrsim 1 \quad (8.4.3)$$

This is nearly the same criterion that I derived from more rigorous and lengthy arguments (Melosh, 1982b). Assuming that lunar multirings form around craters larger than about 300 km in diameter and that the effective lunar lithosphere at the time they formed was about 50 km thick, the estimate for  $S$  presented above indicates that  $Y$  is about 50 MPa, a value in line with other estimates of the strength of silicate planetary lithospheres in extension. More systematic application of Eq. 8.4.3 to lunar multiring basins suggests that the yield stress  $Y$  is in the 40- to 70-MPa range. The strength of ice in extension is only a few tenths of MPa, so that the criterion of Equation 8.4.3 is exceeded many times over for the basins on Callisto or Ganymede, assuming lithospheres a few tens of km thick.

Valhalla-type multiring basins form when the transient crater's depth greatly exceeds the lithosphere's thickness and the lithosphere is disrupted to great distances from the initial crater (Fig. 9.11c). In this case the shear stress  $S$  is so large that the lithosphere quickly fragments into multiple rings, which may be in the form of graben whose width is comparable to the lithosphere thickness. Rather than treating the lithosphere's failure on a fracture-by-fracture basis, it must be regarded as a thin plastic sheet that is characterized by zones of failure. Computations of the extent and character of these zones (Melosh, 1982b) yields results that agree well with the deformation observed around Valhalla and the other Callisto basins.

The fact that the more distant scarps surrounding Valhalla face outward is also explained by ring tectonic theory (see Figure 9.13). Near the rim of the transient crater the asthenosphere flows rapidly inward toward the crater cavity. This rapid flow moves inward relative to the more slowly moving lithosphere and exerts a torque on the lithosphere that tilts the crustal blocks outward, raising inward-facing scarps. More than a few crater diameters farther away, however, the inward flow is weak or even changes sign to become outward as a result of surface loading by the ejecta blanket. At such distances from the crater's center, stresses are mainly transmitted horizontally through the lithosphere. Because the lithosphere's inner edge is pulled inward, so are its more distant portions, although to a lesser degree. Far from the crater's rim the lithosphere thus moves inward relative to the asthenosphere, and the resulting torque tends to tilt crustal blocks inward, raising outward-facing scarps. At intermediate distances simple graben are expected to form, as is observed.

The final crater type predicted by ring tectonic theory (Fig. 9.11d) has not been observed on any planet or satellite to date. This crater type forms when the asthenosphere is so fluid that waves form within it as the crater is excavated. These waves travel outward from the crater, fracturing



**Fig. 9.13** Flow within the interior of a planet subsequent to an impact that penetrates its lithosphere. The presence of a stiff mesosphere beneath a more fluid asthenosphere directs the flow inward toward the crater cavity in the vicinity of the crater rim, producing circular extensional faults (ring scarps) that face inward. Farther from the crater rim a combination of radially outward flow from the subsiding rim and inward transport of the lithosphere produces a torque couple that tends to tilt lithospheric blocks inward, raising outward-facing scarps along the faulted edges. A few graben may form in between these two regimes.

the lithosphere as they go. The initial motion in such a wave is upward and outward, producing an initial radial set of fractures that are cross cut by shorter concentric fractures that form later. Although fractures of this type have not been observed on the surfaces of extraterrestrial bodies, they were produced by the U.S. Army Corps of Engineers during the 1967 "Operation Peggy," which involved high-explosive detonations on floating ice.

The principal factor that decides whether Valhalla-type rings or radial and concentric fractures will form is the viscosity of the underlying asthenosphere. If the asthenosphere is so fluid that the crater floor oscillates up and down, rather than simply rising to near the preimpact surface, outward-moving waves are generated. This criterion is identical to that presented in the last chapter (Eq. 8.3.6), which indicates whether central peaks will occur in a crater. The only difference is that central peaks form in a small region of fluidized debris in the immediate vicinity of the crater, whereas radial and concentric fractures of the type illustrated in Figure 9.11d must be underlain by a broad region of uniform viscosity. Acoustic fluidization associated with crater excavation does not extend far enough from the crater to produce this type of fracture pattern.

The fact that the radial-and-concentric fracture pattern has not been observed on either Ganymede or Callisto implies that neither satellite had a liquid water mantle when the presently observable craters were formed. Liquid water's viscosity is about  $10^{-3}$  Pa-second, whereas the viscosity of Callisto's mantle must have been *at least*  $1.4 \times 10^{12}$  Pa-second to suppress wave formation following the impact that produced the Valhalla structure (by Eq. 8.3.6).

Ring tectonics relates the structural style of multiring basins to the rheology of the planetary surface in which the basin forms. The primary determinants of style are lithosphere thickness and the existence of a low-viscosity channel beneath it. Thus, the scarcity of multiring basins on Mer-

cury and Mars implies that they had thick lithospheres at the time when the largest impacts took place, a conclusion supported by other types of evidence. The moon's lithosphere was relatively thin by comparison (ca. 50 km; see Figure 11.15), while the lithospheres of Ganymede and Callisto were thinner and weaker than the moon's (ca. 10-km thick). The fact that the most ancient impacts on both Ganymede and Callisto produced extensive ring systems whereas more recent impacts of comparably large bodies produced lunar-type ring systems, such as Gilgamesh, or none at all (McKinnon and Melosh, 1980), strongly suggests a secular thickening of the lithosphere due to cooling. This secular trend is similar to that documented by a study of viscous relaxation of smaller craters (Passey and Shoemaker, 1982). The absence of the radial-and-concentric multiring style also places lower limits on the viscosity of the underlying mantle and indicates that none of the solar system bodies with which we are familiar had liquid mantles when their present surfaces formed.

## 9.5 MASCON TECTONICS

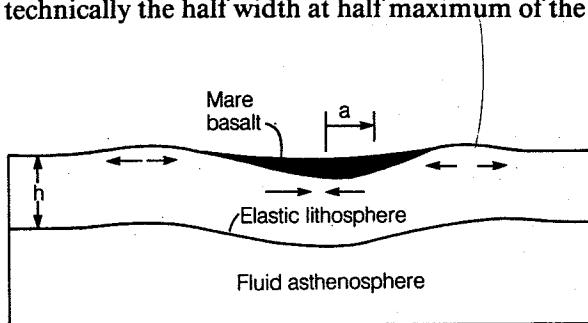
Mascons, an acronym for "mass concentrations," were first discovered in 1968 when large perturbations were noted in the orbits of the Lunar Orbiter spacecraft series. Mascons are anomalous mass excesses associated with nearside lunar basins. After several years of speculation and study it became clear the excess mass is due to the mare basalts filling the basins. Maintained out of isostatic equilibrium by the flexural strength of the moon's lithosphere, mascons are produced by 1 to 3 km of uncompensated mare basalt. Farside basins that lack mare also lack mascons, while Orientale, with only a small central mare, has only a small mascon. Although they are not strictly impact-generated features, they deserve a short treatment in this section because of their common association with impact basins.

While the excess gravitational acceleration pro-

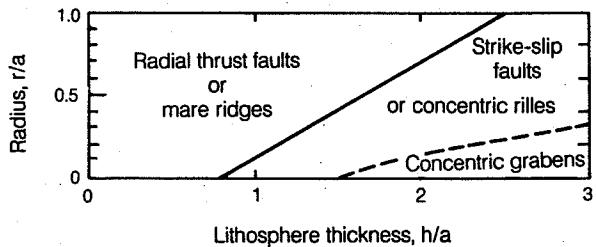
duced by the mascon is of little geologic importance, the tectonic effects of the excess load are profound. The weight of the uncompensated basalt depresses the lithosphere beneath the mare, flexing and deforming both the lithosphere and the pile of mare basalts. Stresses that develop in response to this deformation produce characteristic patterns of faults and folds on the mare surfaces. Since these patterns are diagnostic of lithospheric thickness, knowledge of both tectonic pattern and mascon size allows the lithospheric thickness at the time of mare basalt flooding to be inferred.

Figure 9.14 shows that beneath the mare basalts, stresses in the depressed lithosphere are generally compressive. This compressive stress is transmitted through the pile of mare basalts and may produce compressive tectonic features on the mare surface. Mare ridges, sinuous ridges up to 500 m in height and several hundred kilometers long that lie in crudely concentric patterns on the inner mare surfaces, are believed to result from such stresses. Mare ridges have been variously interpreted as faults (one of them connects with the Lee-Lincoln thrust fault in the Apollo 17 landing area) or as folds. Some evidence even suggests a component of strike-slip motion in addition to compression. These tectonic manifestations are generally consistent with the expected stress field beneath the mascon.

Farther away from the mascon's center, the radial stress near the surface becomes extensional. Either strike-slip faults or concentric normal faults, depending upon depth of initiation (Golombok, 1985), are expected to form in this region. The presence of concentric arcuate rilles, interpreted as grabens, is consistent with this stress field. The boundary between the inner compressive tectonic zone and the outer extensional zone is a sensitive function of lithosphere thickness, as shown in Figure 9.15. Once the mascon radius  $a$  is deduced from its gravitational acceleration ( $a$  is technically the half width at half maximum of the



**Fig. 9.14** The weight of mare basalts filling a basin depresses the lithosphere, producing compressive stresses within the basin and extensional stresses near the surface around it. The mean half-width of the load is  $a$ , and the lithosphere thickness is  $h$ . Magmas intruded directly beneath the basin are guided by the stress field to emerge on the surface near its outer rim.



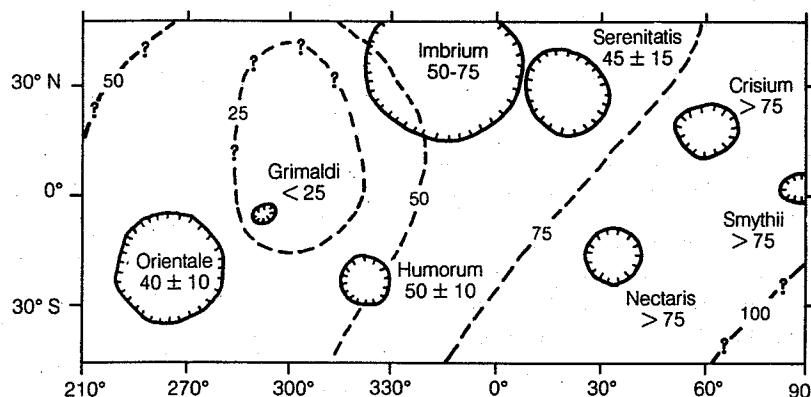
**Fig. 9.15** The boundary between the compressional tectonic regime in the center of a mascon and the extensional regime surrounding it is a sensitive function of lithosphere thickness, here normalized by the mean half-width  $a$  of the mascon load. If  $a$  and the radius of the tectonic transition are known, the thickness of the lithosphere at the time of mare basalt flooding may be read off this plot.

mascon load), the radius of the tectonic transition permits the lithosphere thickness to be read directly off Figure 9.15.

The results of an intensive study of the lunar lithosphere thickness using this technique (Solomon and Head, 1980) are illustrated in Figure 9.16. The lithosphere thicknesses thus derived are consistent with the extent of ring development around large impact craters, as predicted by ring tectonic theory. Grimaldi, for example, is one of the smallest multiring basins on the moon, and it occurs in an area where the elastic lithosphere thickness is smaller than anywhere else on the lunar nearside. Maria Crisium and Smythi occupy large impact structures that lack well-developed exterior rings. The lithosphere is concomitantly thick in this area. The two independent fields of ring tectonics and mascon tectonics thus qualitatively agree in their estimates of lithosphere thickness on the moon.

A mascon of a different kind occurs in the vicinity of the Caloris basin on Mercury. The orientation of Mercury suggests that a massive positive gravity anomaly is associated with Caloris. However, although there is a mare-like lava fill within the basin, the tectonic style in the basin center is indicative of extension, not compression. The positive gravity anomaly in this case seems to be due, not to lavas within the basin, but to the voluminous volcanic smooth plains surrounding the basin. These plains were emplaced shortly after the excavation of Caloris and are probably associated with thermal disturbances caused by the impact. Radar altitude profiles across the smooth plains show that Mercury's lithosphere is sagging down beneath their weight. Caloris, then, is a mascon in which the load is concentrated in an annulus surrounding the impact structure rather than within it, as in lunar mascons. The ultimate origin of the mass anomaly, however, is the same in both: a superisostatic load of flood lavas erupted subsequent to the impact.

## IMPACT CRATERING



**Fig. 9.16** Lithosphere thickness on the lunar nearside. The thickness was determined for each mascon basin from the radius of the transition from compressional to extensional tectonics. Contours are extremely crude. The lithosphere thickness correlates well with the development of multiring basins, as predicted by ring tectonic theory. After Solomon and Head (1980).

## 9.6 THERMAL EVOLUTION OF IMPACT BASINS

The thermal effects of a very large impact on a planetary surface are twofold. First, the impact converts most of the projectile's kinetic energy to heat (see Sec. 5.2.5), much of which is deposited near the site of the impact. Second, collapse of the transient crater lifts temperature contours upward from beneath the crater. The extent of this uplift can be estimated from Equation 8.3.3. The first effect causes a net warming of the target planet, since heat is added. The second effect causes a net cooling of the planet, since it brings hot material from depth closer to the surface where it is conducted more readily to the surface. At the present time it is not clear which effect dominates. Under certain circumstances, such as during the formation of a thin conductive crust over the ancient lunar magma ocean, it seems clear that impacts that disrupt the crust hasten the cooling of the planet's interior. On the other hand, most current models of planetary accretion assume that some fraction of the infalling material's kinetic energy is buried within the growing planet and plays an important role in its differentiation.

Cooling of a planet's lithosphere after a large impact may also have tectonic consequences. Whether the excess heat arises from the kinetic energy of the projectile or from uplift of underlying hotter material, it increases the surface heat flow in the vicinity of the impact and causes a relatively rapid decline in temperature beneath the collapsed crater. The timescale for this decline is on the order of  $h_{SU}/k$ , where  $k$  is the thermal diffusivity (on the order of  $10^{-6}$  m<sup>2</sup>/second for rock) and  $h_{SU}$  is the stratigraphic uplift due to crater collapse (Eq. 8.3.3). The rock beneath the crater contracts as it cools. This contraction causes subsidence of the basin center, which, for the Orientale basin on the moon, may have reached a kilometer. Elastic stresses arise from differential contrac-

tion of the cooling rock, once its temperature drops below an "elastic blocking temperature" at which elastic stresses begin to accumulate (about 800°C). These stresses may have reached several hundred megapascals in the vicinity of Orientale's center and are thought to be responsible for the extensive fissuring on the plains units just inside the Inner Rook mountain ring.

Much work remains to be done before multi-ring basins can be said to be fully understood. They form the major structural element in the lithospheres of the moon, Ganymede and Callisto, and might be important on the earth, Mercury, and Mars, although there is some doubt that they actually do develop on these latter planets. The formation of multiring basins appears to depend upon the rheologic structure of the bodies in which they form. If we can understand these conditions properly, the style of multiring-basin development may be used to probe conditions in the interior of the planet at the time of basin formation.

## FURTHER READING

### General References

The interesting story of the recognition of multi-ring basins, along with an analysis of the reasons why other workers failed to recognize them, is told by

W. K. Hartmann, Discovery of multi-ring basins: Gestalt perception in planetary science, in *Multi-ring Basins, Proc. Lunar and Planet. Sci.*, 12A, pp. 79-80 (1981).

The discovery paper, written by W. K. Hartmann and G. P. Kuiper, was unfortunately published in the limited-circulation *Communications of the Lunar and Planetary Laboratory*, 1, pp. 51-66 (1962). It is not noted as a primary reference here because it is so difficult to obtain. A number of reviews of lunar multiring basins have been writ-

ten. One of the most complete is

W. K. Hartmann and C. A. Wood, Moon: Origin and evolution of multi-ring basins, *The Moon* 3: 3-78 (1971).

This reference also contains a discussion of the famous  $\sqrt{2}$  ring diameter ratio.

### Special Topics

The subsurface structure of lunar multiring basins is the subject of

S. R. Bratt, S. C. Solomon, J. W. Head, and C. H. Thurber, The deep structure of lunar basins: Implications for basin formation and modification, *J. Geophys. Res.* 90: 3049-3064 (1985).

The morphology of Valhalla-type multiring basins is described by

Q. R. Passey and E. M. Shoemaker, Craters and basins on Ganymede and Callisto: Morphological indicators of crustal evolution, in D. Morrison (Ed.), *Satellites of Jupiter*, University of Arizona Press, Tucson, pp. 379-434 (1982).

The occurrence of lunar-type multiring basins on the terrestrial planets is perceptively discussed by

C. A. Wood and J. W. Head, Comparison of impact basins on Mercury, Mars and the Moon, *Proc. Lunar Sci. Conf.* 7th, pp. 3629-3651 (1976).

The various theories of multiring basin formation begin with the volcanic modification hypothesis, most accessibly described in

W. K. Hartmann and F. G. Yale, Mare Orientale and its intriguing basin, *Sky and Telescope* 37: 4-7 (1969).

The megaterrace hypothesis is most fully developed by

J. W. Head, Origin of rings in lunar multi-ringed basins: Evidence from morphology and ring spacing, in D. J. Roddy, R. O. Pepin, and R. B. Merrill (Eds.), *Impact and Explosion Cratering*, Pergamon Press, New York, pp. 563-573 (1977).

The nested crater hypothesis is developed by

C. A. Hodges and D. E. Wilhelms, Formation of lunar basin rings, *Icarus* 34: 294-323 (1978).

Ralph Baldwin has championed the frozen tsun-

ami theory of basin-ring origin. I described this theory in the last chapter because I feel that it best describes peak-ring craters. However, since Baldwin proposed the theory to describe the formation of basin-ring scarps, I duplicate the reference here:

R. B. Baldwin, The tsunami model of the origin of ring structures concentric with large lunar craters, *Phys. Earth Planet. Interiors*, 6: 327-339 (1972).

The theory of ring tectonics was first described briefly by

H. J. Melosh and W. K. McKinnon, The mechanics of ringed basin formation, *Geophys. Res. Lett.* 5: 985-988 (1978).

The theory was elaborated and applied to Valhalla-type structures by

W. B. McKinnon and H. J. Melosh, Evolution of planetary lithospheres: Evidence from multiringed structures on Ganymede and Callisto, *Icarus* 44: 454-471 (1980).

A more mathematical analysis of the Valhalla basin is given by

H. J. Melosh, A simple mechanical model of Valhalla Basin, Callisto, *J. Geophys. Res.*, 87: 1880-1890 (1982).

Many authors, starting with R. B. Baldwin's 1949 book *Face of the Moon* (p. 199), have suggested a connection between mare ridges and subsidence of the mare basins on the moon. An analytic model that demonstrated the connection between tectonic style and lithosphere thickness was first constructed by

H. J. Melosh, The tectonics of mascon loading, *Proc. Lunar Planet. Sci. Conf.* 9<sup>th</sup>, pp. 3513-3525 (1978).

A great deal of work has subsequently been done on mascon tectonics and lunar lithosphere thickness. A good review of this work is

S. C. Solomon and J. W. Head, Lunar mascon basins: Lava filling, tectonics, and evolution of the lithosphere, *Rev. Geophys. Space Phys.* 18: 107-141 (1980).

The thermal evolution of impact basins is still a relatively new subject. A good beginning has been made by

S. R. Bratt, S. C. Solomon, and J. W. Head, The evolution of impact basins: Cooling, subsidence, and thermal stress, *J. Geophys. Res.* 90: 12,415-12,433 (1985).

## X. CRATERED LANDSCAPES

### 10.1 INTRODUCTION

Impact craters have been treated as individual entities in the preceding pages. However, as space-craft images abundantly illustrate, the surfaces of most planets and satellites are scarred by vast numbers of impact craters that range in size from the limit of resolution to a substantial fraction of the planet's or satellite's radius. In some places impact craters are the dominant landform: little of the observed topography can be ascribed to any other process (see Fig. 10.1). Craters on such a surface exhibit degrees of preservation ranging from fresh craters with crisp rims and bright rays to heavily battered or buried craters that may only betray their presence by a broken rim segment or a ragged ring of peaks.

Long-repeated meteorite bombardment of a once-pristine rocky surface also brecciates the underlying rock and covers the surface with ejecta deposits, resulting in a thin "soil" of rock fragments, fractured mineral grains, and bits of impact melt glass. First recognized on the moon during the Surveyor missions, such an impact-produced regolith is characteristic of the surfaces of most airless bodies in the solar system. Regoliths of this type have been observed on the tiny Martian moons Diemos and Phobos. Many brecciated meteorites, such as the gas-rich meteorite Kapoeta, are believed to be samples of asteroidal regoliths. The rate of accumulation of such regoliths is low, amounting to only meters per billion years of exposure, but on small airless bodies that lack competing geologic processes, impact-generated regoliths may grow to depths of tens of meters.

The regolith is a major feature of a planetary surface at the scale of a human explorer or robot rover. The outlines of angular rock outcrops are softened and debris on hillsides creeps slowly downhill, forming characteristic patterns and structures. Many steep lunar slopes are criss-crossed by "terracettes" that superficially resemble the terrestrial hillslope features commonly ascribed to the trampling of cattle. Rocks even on flat surfaces are rolled and overturned every few million years and, as the Apollo astronauts observed, small "fillets" of fine-grained material collect at the bases of protruding rocks. This surprising mobility of the surface layers of small airless bodies is due to the aggregate effect of many small

impacts that continuously, albeit slowly, overturn and mix the upper layers of the regolith by a process aptly termed "gardening."

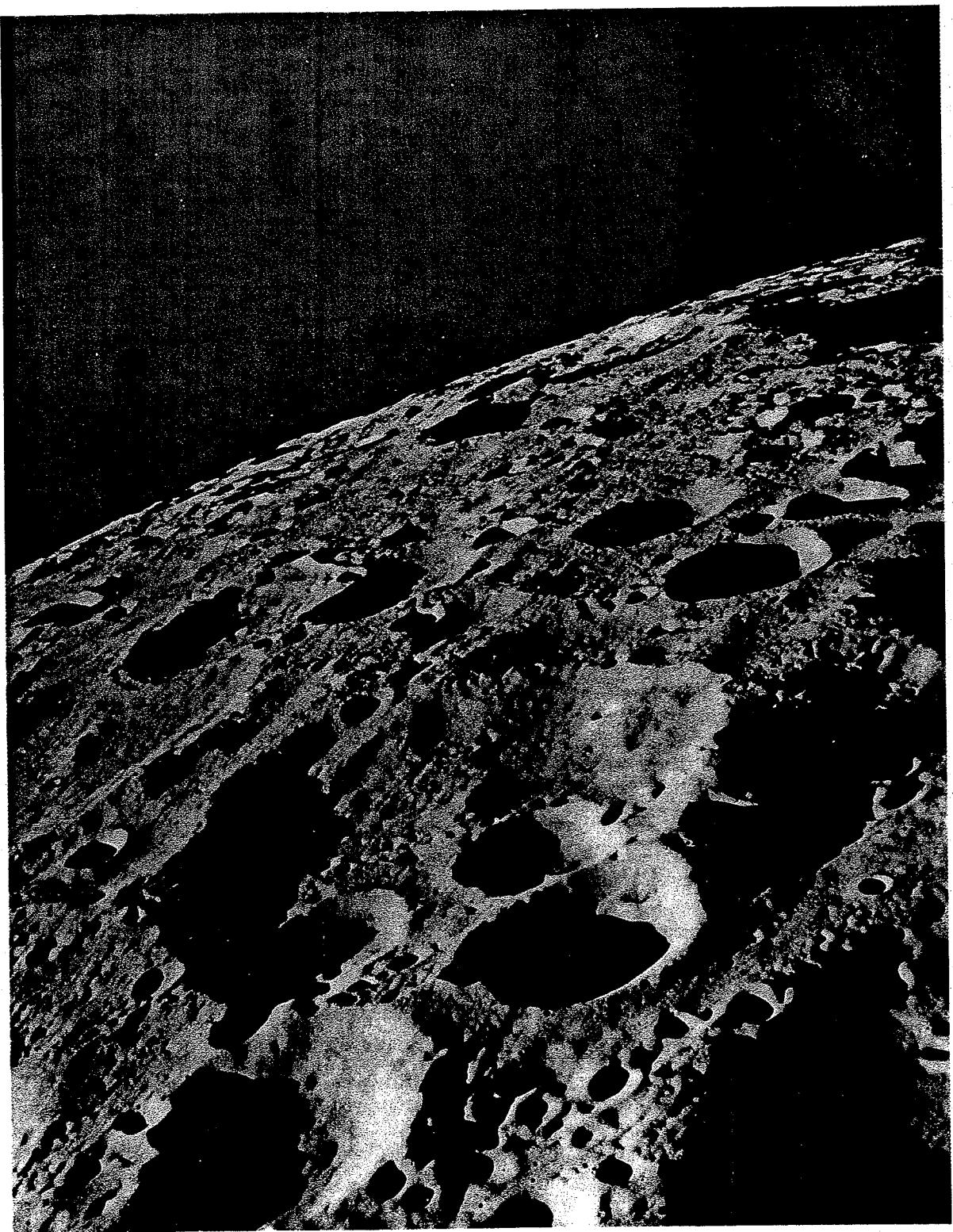
The purpose of this chapter is to briefly acquaint the reader with the specialized techniques for describing the effects of a large number of cratering events on a planet's surface. This is necessarily a mere overview of the large amount of work that has been done in this area, which could itself be expounded at book length. Research on cratered landscapes divides naturally into two parts, reflected in the organization of this chapter. The first part is concerned with the description and evolution of cratered landscapes. Much of this work is derived from efforts to obtain geologic ages for different surfaces in the solar system. The second part is concerned with the nature and development of impact-generated regoliths. This work has been fueled by the return of the lunar samples, most (some lunar geologists would argue all) of which are bits of the regolith. Although these two areas of study have been largely separate, characterized by different needs and methods, repeated meteoritic impacts are fundamental to both. The major difference between them is that cratered landscapes, on the scale that we can commonly observe them, are predominantly due to large impacts (craters 4 km in diameter and larger), whereas regoliths are created and maintained mainly by much smaller impacts.

Because this is a book about impact cratering, the discussion will focus mainly on surfaces whose characteristics are the result exclusively of impact cratering and will thus ignore such other processes as crater obliteration by lava flows, wind, or water. The landscapes developed by such processes are treated at length in the geomorphological literature. The effects of these processes on crater populations, mentioned briefly below, are discussed at length in several of the references given at the end of the chapter.

### 10.2 CRATER POPULATIONS

The present crater population on surfaces such as the lunar highlands (Fig. 10.1) or the more lightly cratered lunar mare is the outcome of a long history of impact cratering events. Analysis of the existing crater population in conjunction with some assumptions about the rate of crater formation

Fi  
Ts  
cr  
15



**Fig. 10.1** Oblique view of a heavily cratered landscape on the moon. This area is to the northeast of Tsilokovsky on the moon's farside. The large crater near the center is about 75 km in diameter, but craters as small as a few tens of meters in diameter can be discerned in the foreground. *Apollo 17 photo 155-23702 (H).*

may reveal a great deal about the geologic history of a surface. A typical population is composed of craters with a wide range of sizes, some of which are relatively fresh, with sharp rims, extensive rays, and crisp fields of small secondary craters, whereas others are progressively more degraded. On parts of such airless bodies as the moon, the principal agent of degradation is other impacts, producing a surface that appears to be crowded with craters. In other regions volcanism has created plains that are more or less sparsely cratered. Elsewhere in the solar system the activities of wind, water, or tectonic processes such as subduction erase craters within a short time of their formation, leading to landscapes like the earth's, where impact craters are among the rarest of landforms, or like that of Mars, where the low abundance of small craters may be due to an ancient era of erosion that obliterated most small craters.

Study of crater populations is thus a powerful tool for geologic investigation of the surfaces of other planets and satellites. If the flux and size distribution of the impacting bodies were known, studies of crater populations could yield absolute ages of the surface and some of its features. Although the original flux is often unknown, relative ages can usually be obtained. Before geologic inferences can be drawn from crater populations, however, we must have an effective means of describing and comparing them. Unfortunately, a large number of descriptions have evolved over the years as each group of scientists studying a particular problem created their own specialized means of presenting population data, making it difficult to compare the results of different groups. In the following section I adopt the major recommendations of a NASA panel convened in 1978 to standardize the presentation of crater population data.

### 10.2.1 Description of crater populations

The first step in an investigation of the crater population on a given surface is to select an area that is believed to have had a homogeneous geologic history. It would make little sense, for example, to combine the crater population of a sparsely cratered lava plain with that of a densely cratered upland. Once such an area is selected, the craters that lie within it are counted. Most crater population studies include all recognizable craters, regardless of their state of degradation. Where a large enough population exists, more specialized studies may be performed in which the numbers of fresh craters, slightly degraded craters, degraded craters, etc., are counted separately. Although these studies leave some room for interpretation as to what a "recognizable" or "fresh" crater is, intercomparisons of results between different groups of crater counters have generally shown good agreement.

Craters occur in a wide variety of sizes, so that the principal information about a crater population is the number of craters per unit area as a function of crater diameter. It is presumed that impact cratering is a random process and that there is no significance to the particular location of craters within the selected area, so that only data on the number and diameter of craters is kept.

Numerous ways of representing the number of craters as a function of diameter have been developed. One very simple method is to list the number  $N$  of craters per unit area with diameters between two limits, say between  $D_a$  and  $D_b$ . The problem with this method is that the resulting number of craters depends upon the interval  $\Delta D = D_b - D_a$ , and different crater counters may choose different intervals. Furthermore, if the interval  $\Delta D$  is fixed at, say, 1 km, this might be convenient for craters with diameters between 5 and 20 km, but would be too large for craters with diameters less than 1 km and too small for craters larger than 100 km. A simple way to overcome this problem is to let the interval depend on crater size. Thus, the number of craters may be tabulated between  $D$  and  $2D$ , where the intervals increase in octaves. Actually, this binning has been found too coarse in practice, so that most such *incremental* size-frequency distributions use an interval of  $D$  to  $\sqrt{2}D$ . The incremental distribution still suffers, however, from the arbitrary choice of a starting diameter  $D$ . It is now recommended that the bins be chosen so that one bin boundary is at  $D = 1$  km.

Although the incremental size-frequency distribution could be successful if the same bin sizes and boundaries are universally adopted, it lacks fundamental simplicity. Another distribution has long been used that is independent of bin size: this is the *cumulative* size-frequency distribution. In this distribution the number  $N_{\text{cum}}$  of craters per unit area with diameters greater than or equal to a given diameter  $D$  is tabulated. Not only is the resulting distribution  $N_{\text{cum}}(D)$  independent of bin size, but any desired incremental distribution can be easily generated from it, since the number of craters  $N$  per unit area in the interval between  $D_a$  and  $D_b$  is simply

$$N(D_a, D_b) = N_{\text{cum}}(D_a) - N_{\text{cum}}(D_b) \quad (10.2.1)$$

where  $D_b > D_a$ .  $N(D_a, D_b)$  is necessarily positive or zero by the definition of the cumulative number distribution. The only disadvantage of the cumulative distribution is that the cumulative number of craters at some given diameter depends upon the number of craters at all larger diameters. Although this is rarely a major problem, cumulative distributions in limited-diameter intervals (often controlled by the size of the region being analyzed) have to be adjusted in overall value to join with the distributions from other diameter ranges.

The sl  
and  
It h  
tive r  
power

where  
betwe  
amete  
also  
earth

It i  
powe  
numt  
dime  
area)

in a c  
such  
tions  
such  
phot  
clues  
abou  
scrib  
a sin  
no fi  
inde  
Equa  
vent  
plot  
slope  
rnull  
have  
diffe  
crate  
desi  
sent  
ence  
the  
that  
use.

T  
pro  
esse  
trib  
ter  
wou  
tior  
bas  
terv  
of t  
cur

$\sqrt{2}$

The slope of the cumulative plot is not, of course, affected by such adjustments.

It has been found in practice that the cumulative number distribution closely approximates a power function of diameter,

$$N_{\text{cum}} = cD^{-b} \quad (10.2.2)$$

where  $b \approx 1.8$  for postmare craters on the moon between 4 and several hundred kilometers in diameter (and, given the limited data, this power also seems to hold for impact craters on the earth).

It is intriguing that the power  $b$  is close to 2. A power 2 in Equation 10.2.2 is a kind of "magic" number because when  $b = 2$  the coefficient  $c$  is dimensionless (remember  $N_{\text{cum}}$  is number per unit area). There is no fundamental length or size scale in a crater population with this power law, so that such a population looks the same at all resolutions. It is impossible to tell from a photograph of such a cratered landscape whether the scale of the photograph is 100 km or 1 m (of course, other clues than crater population alone may give a hint about the actual scale). A population of craters described by a power near 2 might arise either from a simple formation process in which there is truly no fundamental length scale or from a series of independent processes that are so complex and chaotic that no one scale dominates.

Because the cumulative number distribution of Equation 10.2.2 falls as a power of  $D$ , it is conventional to graph such distributions on a log-log plot on which a power law is a straight line with slope equal to  $-b$  (Fig. 10.2). Unfortunately, cumulative number distributions plotted in this way have a tendency to look all the same, apparently differing only in the absolute number density of craters. Although this may be adequate or even desirable for some purposes, such a mode of presentation obscures slight but significant differences in crater populations. Since these differences can be useful in deciphering the source of the crater population or other geologic processes that acted on it, another type of plot is in common use.

This second type of plot exploits the close approach of  $b$  to 2 in Equation 10.2.2 by graphing essentially the ratio between the actual crater distribution and a distribution with slope  $-2$ . A crater population in which the actual slope is  $-2$  would thus plot as a horizontal line. The conventional plot of this type is called an  $R$  plot. It is based on an incremental distribution with  $\sqrt{2}$  intervals between diameter bins. Note that the slope of this type of incremental plot is the same as the cumulative distribution of Equation 10.2.2, although the coefficient is different. It is easy to show that if  $b$  is constant over the interval  $D$  to  $\sqrt{2}D$  then the incremental number density is

$$N(D, \sqrt{2}D) = c(1 - \sqrt{2}^{-b}) D^{-b} \quad (10.2.3)$$

The definition of the  $R$  plot includes several numerical factors. In terms of the cumulative number distribution it is given by

$$R(D) \equiv [2^{3/4}/(\sqrt{2} - 1)]D^2[N_{\text{cum}}(D) - N_{\text{cum}}(\sqrt{2}D)] \quad (10.2.4)$$

Figure 10.2 compares the cumulative and  $R$  plots for post-mare lunar craters.

A useful interpretation of  $R$  is to note that, up to a constant factor,  $R$  is equal to the fraction  $f_c(D)$  of the total area covered by craters in the diameter interval  $D$  to  $\sqrt{2}D$ . Since  $f_c$  is just the geometric mean area of a crater in this interval,  $\pi\sqrt{2}D^2/4$ , times the number in the interval,  $N(D, \sqrt{2}D)$ ,

$$f_c = (\pi\sqrt{2}/4)D^2[N_{\text{cum}}(D) - N_{\text{cum}}(\sqrt{2}D)] \quad (10.2.5)$$

Comparison of definitions in Equations 10.2.4 and 10.2.5 yields the simple relation

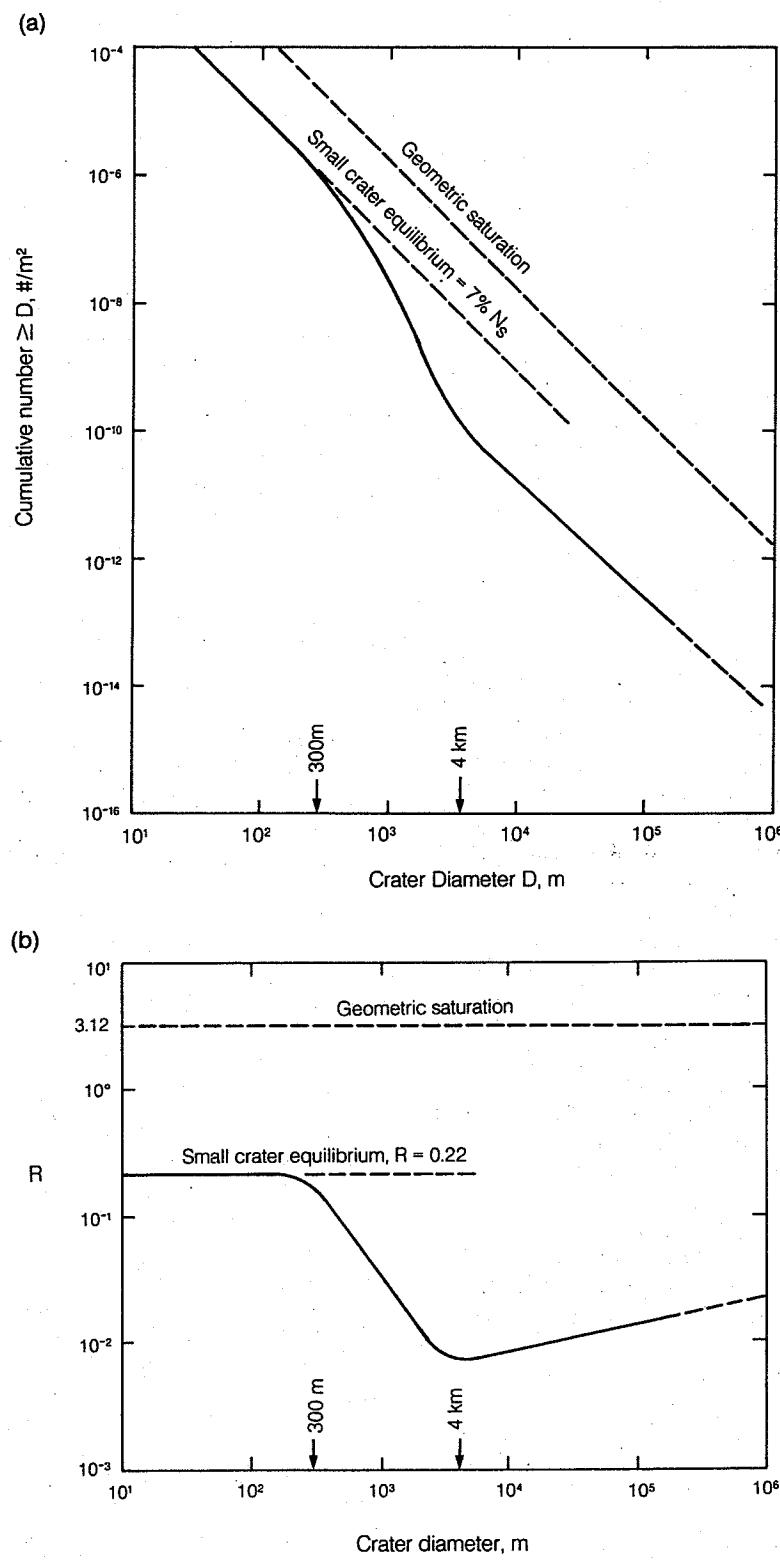
$$R(D) = 3.65f_c(D) \quad (10.2.6)$$

where both  $R$  and  $f_c$  are dimensionless numbers. With this interpretation it is easy to see that in a crater population with  $b = 2$ , for which  $R$  and  $f_c$  are constant, craters in every size interval occupy the same fraction of the total area. If  $b < 2$ , as it is for postmare craters on the moon,  $R(D)$  increases as  $D$  increases so that large craters occupy a larger fraction of the surface than small craters. If  $b > 2$ , small craters occupy a larger fraction of the surface than do large ones.

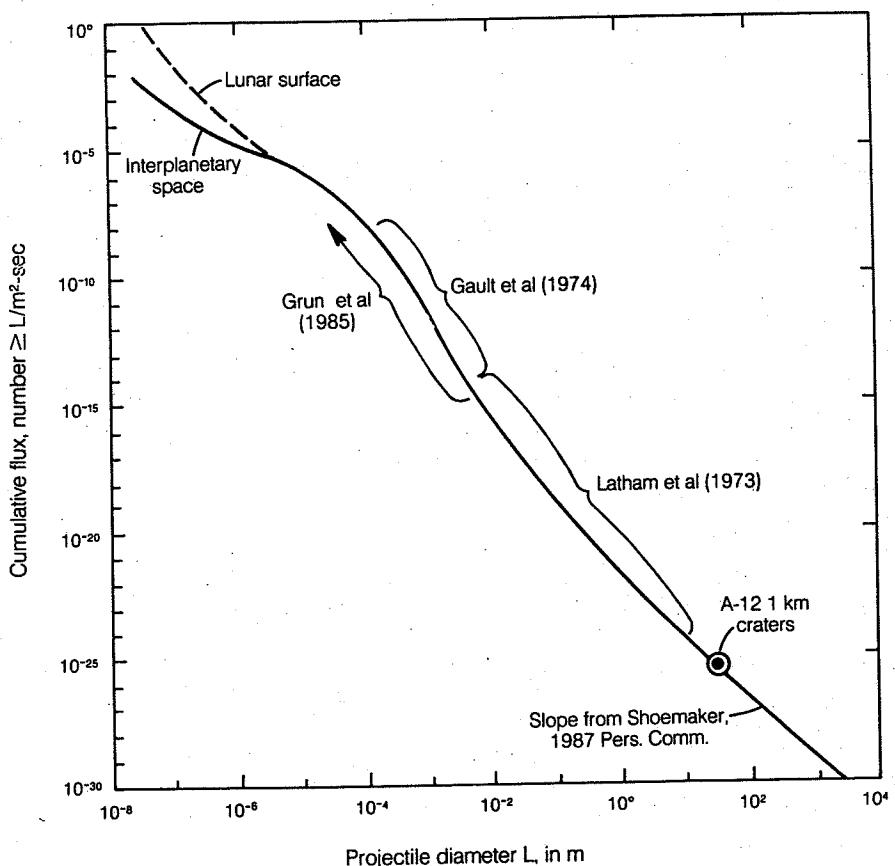
### 10.2.2 The primary flux

The crater population on a planetary surface is ultimately controlled by the flux of meteoroids (including comets) striking it. The connection between the flux and crater population, however, is not direct. As discussed in Chapter 7, the size of the final crater is a function of the velocity and angle of impact, the composition of the projectile and target, and many other factors. The final crater size is also affected by postimpact collapse and enlargement, discussed in Chapter 8. Furthermore, the crater population on a planetary surface is generally modified by the formation of secondary craters around large primaries and by crater obliteration processes such as mutual overlap, burial by lava flows, or erosion. In spite of these caveats, however, it is useful to look at what is known of the flux of impacting bodies in the earth-moon system, illustrated in Figure 10.3.

The data presented in this figure is derived from a variety of sources, ranging from measurements of impacts on spacecraft at the smaller sizes, through a seismic study of impacts on the moon and of terrestrial meteors at intermediate sizes, to direct telescopic observation of earth-crossing asteroids and comets at the largest sizes. The major feature of this plot is an apparent



**Fig. 10.2** The post-mare crater population on the moon on both a cumulative plot (a) and an  $R$ -plot (b) format. The population shows three distinct segments. At small diameters ( $D < 300$  m) the population is in equilibrium and the cumulative number is proportional to  $D^{-2}$ .  $R$  is constant for this population. At intermediate diameters ( $300 \text{ m} < D < 4 \text{ km}$ ) the cumulative number is proportional to  $D^{-3.4}$ , whereas at larger diameters ( $D > 4 \text{ km}$ ) it falls as  $D^{-1.8}$ . The curves are dashed at the largest diameters ( $D > 200 \text{ km}$ ) because no craters of this size have yet formed on the mare. The changes in slope of the crater population are especially evident in the  $R$ -plot format. The dashed line labeled "geometric saturation" is an upper limit to the crater density on any surface. Data from Basaltic Volcanism Study Project (1981), page 1114, plot 11.



**Fig. 10.3** The flux of crater-producing projectiles in the vicinity of the earth and moon. The solid curve represents the flux in interplanetary space, and the dashed curve shows the flux on the lunar surface. These two curves differ only for micrometeorites which are more abundant on the moon because many small high-velocity fragments are produced as secondaries. Note the change in slope of the main curve at a projectile diameter of about 10 m. This change in slope may correlate with the change in slope of the small crater population near 4 km in Figure 10.2. *Data from sources cited in the figure.*

change in slope of the size-frequency distribution of the primary impactors. Leaving out the decrease in slope at the very smallest dust sizes (which is due to processes that clear small particles out of interplanetary space), it appears that the cumulative number of impactors falls roughly as  $L^{-3.5}$ , where  $L$  is projectile diameter, up to diameters of about 10 m. The slope then decreases to roughly  $L^{-2}$  at diameters between 10 m to about 10 km, the largest size yet observed. The cause of this change in slope is presently unknown, although it may be related to the collisional processes in the asteroid belt that produce the smaller-diameter meteoroids.

Over a diameter range from about 100  $\mu\text{m}$  to about 10 m, the cumulative flux  $F_{\text{cum}}$  of impactors on the moon's surface is

$$F_{\text{cum}} = 5.6 \times 10^{-23} L^{-3.48}, \quad 10^{-4} < L < 10 \text{ m} \quad (10.2.7a)$$

where  $F_{\text{cum}}$  is in units of number per square meter per second and projectile diameter  $L$  is in meters. This equation is derived from the more com-

monly quoted distribution in terms of projectile mass  $m$  (in kg)

$$F_{\text{cum}} = 2.5 \times 10^{-19} m^{-1.16}, \quad 10^{-10} < m < 10^5 \text{ kg} \quad (10.2.7b)$$

Spherical projectiles with density 2700 kg/m<sup>3</sup> are assumed in the derivation of Equation 10.2.7a from Equation 10.2.7b.

The flux of projectiles larger than about 1 m in diameter is not well known, but is roughly given by

$$F_{\text{cum}} \approx 1 \times 10^{-23} L^{-2}, \quad L > 10 \text{ m} \quad (10.2.8)$$

When this data is translated into the expected crater population on the moon's surface, it suggests that the slope of the primary crater size-frequency distribution changes in the vicinity of 100-m diameter craters (roughly the size crater produced by a 10-m diameter projectile striking at a few tens of km/sec). Indeed, Figure 10.2b shows just such a change, where  $R$  decreases with increasing diameter for  $D \lesssim 100$  m, indicating  $b > 2$  in this range, whereas  $R$  increases with increas-

ing diameter for  $D \gtrsim 100$  m, indicating that  $b < 2$  in this range. The post-mare crater population on the moon thus seems to reflect the size-frequency distribution of known impactors at the present day. It has also been suggested that the change in slope of the post-mare crater population is caused by secondary craters from large distant primaries, but since the slope break also seems to occur in the primary projectile population, secondaries are evidently not the sole reason for the slope change.

The flux of the projectile population is also in reasonably good agreement with the observed cratering rate of about  $2 \times 10^{-14}$  craters with  $D > 4$  km/km<sup>2</sup>/year on the moon and  $1.8 \times 10^{-15}$  craters with  $D > 22.6$  km/km<sup>2</sup>/year on the earth. Although there are many uncertainties in these figures, there seems no reason at the present time to suspect any major inconsistency between the population of post-mare craters and objects in space.

Although it goes somewhat beyond our need in this chapter, I note that there is strong evidence that the pre-mare crater and hence projectile population differed substantially from the present population. Not only was the population different, but the cratering flux was many orders of magnitude higher than the present flux during the era of "late heavy bombardment" before about  $3.8 \times 10^9$  years ago. A more speculative suggestion has recently been made that the cratering flux on the earth and moon has varied periodically in the post-lunar-mare era, being constant only in an average sense. Events such as hypothetical comet "showers" and the rare breakup of large objects in the asteroid belt may produce such variations. Much work needs to be done before the reality of such events is established, however.

### 10.2.3 The production population

The impact of the primary meteoroid flux on a planetary surface results in some definite rate of crater production as a function of diameter. As the surface, initially taken to be craterless, ages, more and more craters accumulate on it. The integral of the crater production rate over the age of the surface is a special, theoretical, crater population called the *production population*. The production population is the size-frequency distribution of all the craters, excluding secondaries, that have ever formed since craters began to accumulate on the surface. The population is theoretical in the sense that it neglects all crater obliteration processes and is hence formally unobservable, although the crater population on lightly cratered surfaces may approach the production population closely enough for practical applications.

The production population is a useful concept for the study of the evolution of crater populations. Such studies usually begin with an assumed

or inferred production population and then postulate a model of crater obliteration that will, it is hoped, result in a predicted crater population that matches the observed population. The crater obliteration model is a function of the process being modeled. In the next section we shall consider only the process of crater obliteration by other craters. As we shall see, this can be a surprisingly complex process.

### 10.2.4 Impact evolution of crater populations

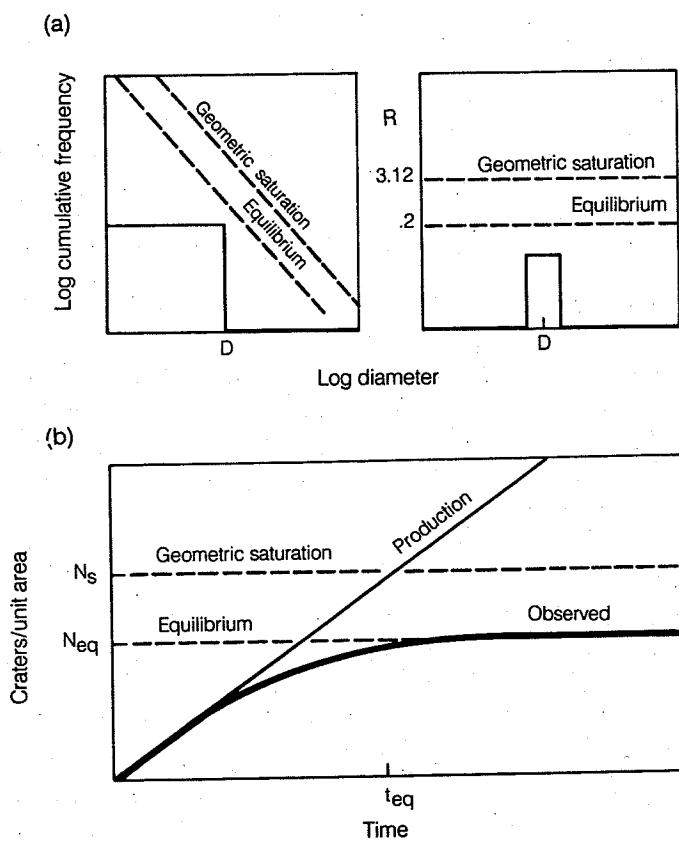
Given the fact that craters accumulate randomly on the surface of a planet at a rate that is, on average, constant, it should be possible to begin with a knowledge of the rate at which craters form as a function of their size and then predict the crater population at any future time. Furthermore, it should be possible, within certain limits, to invert an evolved crater population and deduce both how long it has been accumulating (that is, the age of the surface) and the rate of crater formation as a function of size.

Simple as these propositions may seem, there has been great difficulty in actually implementing them. The main problem is the interaction between craters of different size: the formation of a single large crater on a surface may obliterate many smaller craters, while it takes many small craters to batter a large crater beyond recognition. It has taken many years to fully understand the effects of this interaction on crater populations, and some aspects of it are not completely understood today.

The conceptually simplest population is one in which all craters are the same diameter. Although no natural examples of such a population are known, study of its evolution introduces several important concepts. Moreover, there are crater populations, such as that on Mimas, in which many of the craters fall within a relatively narrow size range and that may thus be approximated by a population of craters all of the same size.

Figure 10.4a illustrates the form of the cumulative and  $R$  plots for a population of craters all of the same diameter. Note that the  $R$  plot does not show an infinite sharp peak, even though it is an incremental distribution, because the bin size is defined to be a finite interval. The cumulative crater distribution is a step function, with the step occurring at the crater diameter  $D$ .

The evolution of this crater population is illustrated in Figure 10.4b as a function of time. The plot shows that the observed crater population and the production population initially increase at the same rate. However, as the density of craters on the surface increases a few older craters are either overlapped by new ones or are buried by their ejecta. As this process continues, some older craters are completely obliterated by younger craters and the observed crater density falls below



**Fig. 10.4** Evolution of a crater population in which all craters are the same size. Part (a) illustrates the cumulative and R-plots for this very special kind of population. Part (b) shows how the crater density increases as a function of time. Although the production population rises linearly with time, the number of craters that can be counted on the surface eventually reaches a limit well below the geometric saturation limit. Once the population has reached equilibrium, each additional crater obliterates, on average, one old crater.

the production line. Eventually, the crater density becomes so high that each new crater that forms obliterates, on average, one older crater. At this stage the crater population has reached *equilibrium*: no further increase in crater density is possible, although new craters continue to form and the production population increases steadily in number.

A note on terminology is necessary here. I have defined the term *equilibrium* (of a crater population) following D. E. Gault's (1970) original definition. Many recent authors, however, use the term *saturation* for the same concept and reserve *equilibrium* for the more general case where craters are obliterated by other geologic processes in addition to impact. This terminology is especially confusing because Gault and several others during the early 1970s used *saturation* for a somewhat different concept, one that I shall introduce shortly as *geometric saturation*. Several other variants of the concept of *saturation* occur in the literature, about which silly but furious battles have been (and are being) fought. I have chosen to return to Gault's original term here but the

reader may, of course, make up his or her own mind as to whether the term is apt.

The attainment of equilibrium places severe constraints on attempts to date planetary surfaces by crater counting. Up until equilibrium is attained at time  $t_{eq}$  in Figure 10.4 the crater density increases with the age of the surface, so that knowledge of the crater production rate permits computation of an absolute age from the crater density. Even if the production rate is not known, the relative ages of two surfaces may be obtained by comparing their crater densities. However, once equilibrium is attained the crater density becomes constant and only a lower limit on the age can be obtained. The relative ages of two surfaces in equilibrium are also completely unconstrained, since their crater densities are identical even though the surfaces may be widely different in age.

A second useful concept introduced by Gault will be called here *geometric saturation*, following the authors of NASA's 1981 *Basaltic Volcanism Study Project*. Gault introduced the idea in an attempt to quantify the absolute upper limit to the

crater density on a planetary surface. He also supposed (we now know incorrectly in some cases) that the equilibrium crater density is a simple fraction of this maximum possible crater density. The number density  $N_s$  of a hypothetical "geometrically saturated" crater population is equal to the number of craters of diameter  $D$  that can be placed on the area fitted rim to rim in hexagonal closest packing. Since this is to be an upper limit to the possible crater density, no provision is made for obliteration of adjacent craters by ejecta blankets, nor even for random spacing. It is easy to show that for a population of craters of the same size  $D$ ,

$$N_s = 1.15D^{-2} \quad (10.2.9)$$

Another way of stating the same result is that at this density craters occupy 90.5 percent of the total area in hexagonal closest packing. Because real craters are randomly distributed, the production crater density must exceed about  $2.6N_s$  before an actual coverage of 90.5 percent is obtained. At this density the surface is truly "saturated" in the sense of saturation bombing, where nearly every part of the surface has been struck by at least one crater.

Although the study of populations of equal-sized craters has led to some useful concepts, crater populations in nature are composed of craters with a range of sizes. Generalizing the concepts of geometric saturation and, especially, equilibrium to such populations has turned out to be more difficult than it appeared to be at first.

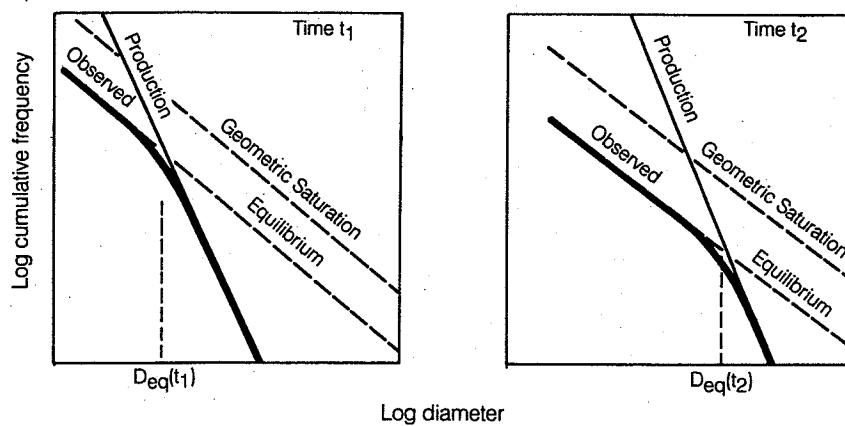
The concept of geometric saturation is less easy to define over a range of crater sizes because the cumulative number of craters depends on the diameter interval over which hexagonal closest packing is supposed to occur. Clearly, the cumulative crater density would be infinite if craters

were distributed in hexagonally closest-packed arrays within each infinitesimal diameter increment. To make some sense out of the concept for a range of crater sizes, Gault supposed that craters that differ by a factor 2 in diameter could be considered to belong to distinguishable packed crater arrays, so that the cumulative saturated crater density,  $N_{cs}(D)$  is

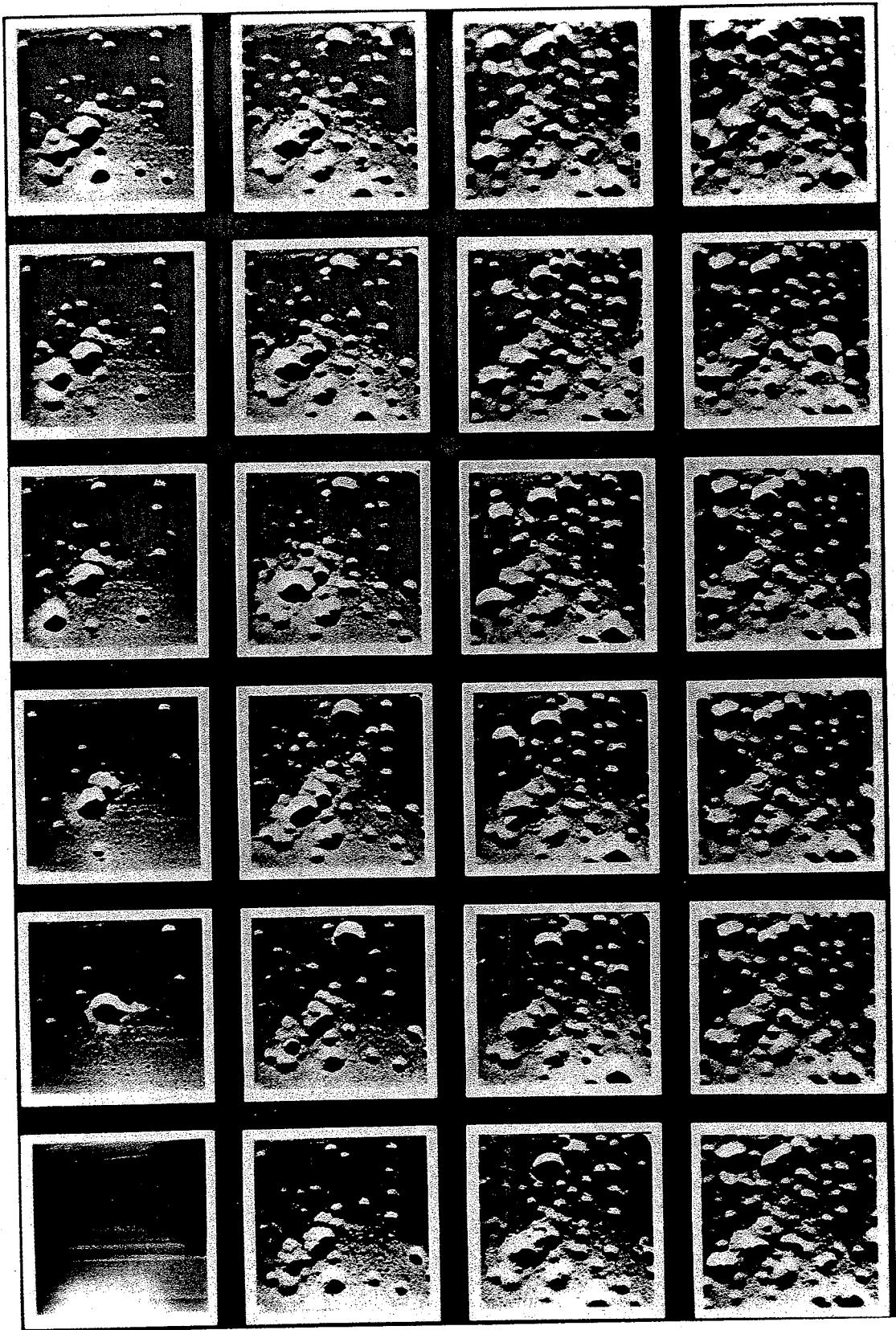
$$N_{cs}(D) = 1.54D^{-2} \quad (10.2.10)$$

This density corresponds to  $R = 3.12$ , or a fractional surface area coverage (Eq. 10.2.6) in each  $\sqrt{2}$  diameter interval of  $f_c = 0.85$ . The geometric saturation lines defined by these equations are plotted for reference purposes in Figure 10.4a. Remember that geometric saturation is supposed to be a theoretical limit that can never be attained by any real crater population. The maximum density measured to date is on Mimas, where  $R$  reaches about 0.4 (13 percent of geometric saturation) for craters between 10 and 20 km in diameter, well below the limit. A number of simulations performed by Gault (which used a range of crater diameters) reached equilibrium at crater densities between 1 percent and 10 percent of the geometric saturation density, with 5 percent to 7 percent being a preferred mean.

The attainment of equilibrium by crater populations divides into two distinct cases (and a trivial intermediate). The first, and simplest, case occurs when the production population has a slope  $b$  steeper than 2. This case was studied by Gault and offers the fewest conceptual difficulties. Unfortunately, this distribution is only appropriate for small ( $\leq 100$  m) craters on the lunar mare (see Figure 10.2), although at the time Gault performed his analysis it was believed to be valid for all crater sizes. The size-frequency distribution of larger lunar craters follows a power law with a



**Fig. 10.5** Evolution of a crater population with slope  $b > 2$ . The production population exceeds the equilibrium line at small crater diameters. Small craters are thus in equilibrium up to some diameter  $D_{eq}$ , above which the observed population follows the production population. The left panel illustrates the population at a relatively early time  $t_1$  and the right panel shows how the population has changed at a later time  $t_2$ . The equilibrium diameter  $D_{eq}$  clearly increases as a function of time, although this increase is generally not linear.



**Fig. 10.6** A laboratory scale demonstration of the concept of crater equilibrium. The photographs are of a box 2.5 m square filled 30-cm deep with quartz sand. The sand is topped with 2 cm of carbon-dum powder to provide a color contrast. Six sizes of projectile were fired into the box at random locations, simulating a production population with slope index  $b = 3.3$ , similar to that of small craters on the moon. Time increases from upper left horizontally to lower right. Equilibrium is attained about halfway through the simulation: although individual surface details vary from frame to frame, the crater population in the later frames remains the same. *From Gault (1970); photo courtesy of R. Greeley.*

slope  $b \approx 1.8$ . This distribution fits the second case of an evolving population with slope  $b$  smaller than 2, and will be treated shortly.

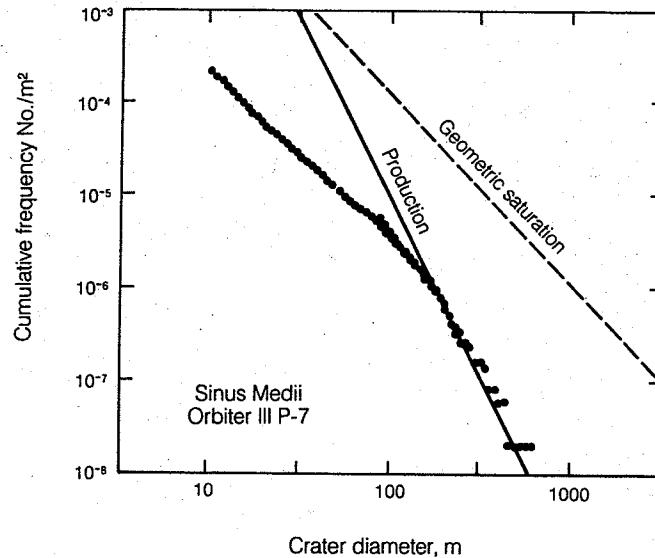
Figure 10.5 illustrates the evolution of a production population with slope  $b$  steeper than 2. The left frame depicts the population at an early time  $t_1$  and the right frame is at a later time  $t_2$ . Because the production population is steeper than the geometric saturation line, mutual obliteration must occur for sufficiently small craters (craters smaller than  $D_e(1)$  in Figure 10.5) no matter how early the time (unless, of course, the population has had so little time to evolve that the statistics of small numbers of craters begins to play a role). Since the actual crater density cannot reach the geometric saturation limit, the observable crater density reaches equilibrium somewhere below this line. It seems intuitively reasonable that the equilibrium population should follow a line parallel to, but below, the geometric saturation distribution. Gault showed empirically in small-scale impact experiments (Fig. 10.6) that this is the case, and subsequent work has confirmed this result both analytically (Soderblom, 1970) and by Monte Carlo computer simulations (Woronow, 1977). In Gault's experiments the crater equilibrium density depends on the slope  $b$  of the production population, with steeper slopes giving a lower equilibrium crater density. Similar results were also obtained from the theoretical studies.

The observed population at any one time in Figure 10.5 is thus composed of two branches. Small craters follow an equilibrium line with slope  $b = 2$ . Larger craters follow the steeper production population curve. The inflection point

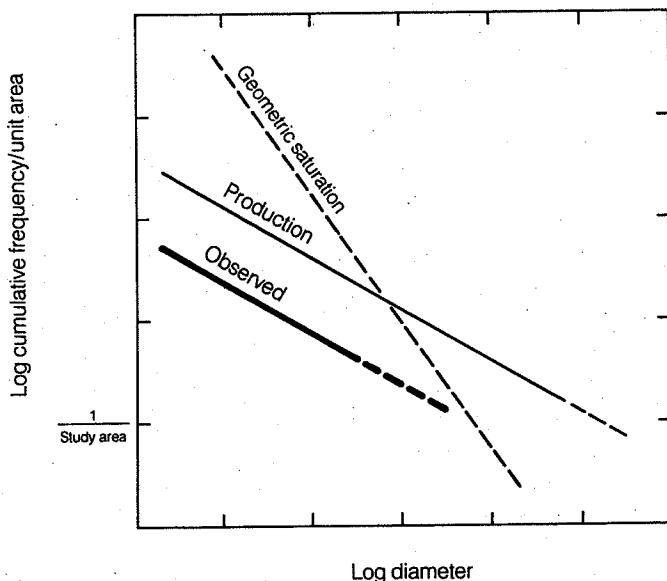
between these two curves is at diameter  $D_e(1)$  where the production curve crosses the equilibrium line. The right-hand frame in Figure 10.5 shows the crater density at a later time. The crater population is qualitatively similar to that at the early time, except that the transition diameter  $D_e(2)$  is larger. If the rate at which craters accumulate is constant (that is,  $c$  in Equation 10.2.2 is a linear function of time), it is easy to show from Equation (10.3.2) that the transition diameter  $D_e(t)$  grows as  $(\text{time})^{1/(b-2)}$ . Conversely, if the crater production rate, slope  $b$ , and the present  $D_e$  are known, the age of the surface can be computed.

The difference between the production line and the observed equilibrium line in Figure 10.5 is the cumulative number of craters that have been obliterated. Because the production line is steeper than the equilibrium line, many more small craters are obliterated than large ones. Essentially no craters larger than  $D_e(t)$  are obliterated at any given time  $t$ . This relation gives rise to the gardening phenomenon discussed more fully in the next section, in which the rate of impact-induced overturn and mixing of an exposed surface decreases rapidly as the depth increases because small craters penetrate only to shallow depths whereas stirring to a greater depth requires a larger, rarer, impact.

Figure 10.7 is a cumulative crater plot for small craters in northern Sinus Medii on the moon. The break in slope between the equilibrium curve and the production population is clearly visible at a diameter  $D_{eq} \approx 100$  m. The production slope  $b \approx 4$  is observable for larger craters. The equilibrium curve in this location is about 3 percent of the



**Fig. 10.7** The small crater population on the lunar surface in the vicinity of Sinus Medii. The observed population clearly shows the attainment of equilibrium at small diameters, whereas the production population, with slope  $b \approx 3.4$ , is visible at large diameters. In this area,  $D_{eq} \approx 100$  m. After Gault (1970).



**Fig. 10.8** Evolution of a crater population with slope  $b < 2$ . The production population exceeds the geometric saturation line at large diameters. Under these circumstances the evolution of the population is dominated by large impacts that allow a number of small craters to accumulate before wiping the surface clean. Under these circumstances the observed population line is always parallel to the production population, but may lie considerably below it. The production population line is dashed at the large crater end because there must be at least one crater within the study area for the line to be meaningful. The observed population may approach geometric saturation if the entire study area is the site of one large impact, but the population will be dominated by fluctuations because of the statistics of small numbers.

geometric saturation limit, in line with other measurements of small crater populations on the moon.

Figure 10.8 illustrates the evolution of a crater population with slope  $b$  less than 2. Note that in the unlikely event that  $b$  exactly equals 2 the observed crater density simply maintains the slope of 2 until the density reaches equilibrium (which occurs simultaneously at all diameters), after which the observed crater density remains constant and the slope remains 2. When  $b$  is less than 2 the situation is more complex. In this case the production curve exceeds the geometric saturation line at the *large* diameter end of the scale. Gault's 1970 model could not deal with this situation and it took nearly 15 years before the implications of this large diameter crossing were understood, in spite of the developing knowledge that  $b \approx 1.8$  for craters more than about 4 km in diameter on the lunar mare. The situation was resolved by C. R. Chapman (Chapman and McKinnon, 1986) who performed a Monte Carlo simulation of crater population evolution that included a wider range of diameters than had previously been possible.

Chapman realized that, first of all, the large-diameter end of the cratering curve is dominated by the statistics of small numbers. Even though the production curve apparently crosses the geometric saturation line at all times, enough time must

pass after the formation of the initial uncratered surface that at least *one* large crater has formed on it. The crossing of the two curves makes little sense unless the cumulative number of craters larger than  $D_s$  is at least 1 in the finite area under study. Since the probability of a small impact is larger than the probability of a large impact for any  $b > 0$ , a population of small craters initially develops that follows the production curve closely. Eventually, however, a large impact occurs. With  $b < 2$ , there is a high probability that this large crater obliterates all or a significant fraction of the study area, wiping out nearly all previous smaller craters. The number of observed craters thus suddenly drops below the production curve. As time passes, small craters again accumulate on the surface. The slope of this new population is equal to that of the production curve but, as shown in Figure 10.7, the cumulative number of such craters is smaller than the production population. The number of small craters continues to grow until the next large impact wipes the slate clean once more.

Under these circumstances there is no "equilibrium" population, although the observed crater density is always well below the geometric saturation limit. The observed crater density fluctuates widely and irregularly, controlled by the large, rare catastrophic impact events. In spite of these wide variations in density, the slope of the

observed population at any given time is roughly equal to that of the production population. The crater densities on a surface of this type are spatially patchy, being low at the sites of recent large impacts and high in areas that have not been struck for a long time by one of the large impacts.

Dating such a surface is nearly impossible after the first large impact, unless some area can be found that has escaped all large impacts. In practice, all that can be determined from the crater density on such a surface is the date since the last large impact that affected the particular study region.

This insight into the evolution of crater populations with  $b < 2$  has resolved an acrimonious controversy over whether the lunar highlands and other heavily cratered surfaces in the solar system are truly "saturated" (in the terminology of the contestants). Although it appears obvious to the casual onlooker that surfaces such as that in Figure 10.1 are so thickly covered with craters no vestige of an original surface could have survived, careful studies using the  $R$  plot have shown that the production population is still discernible. This fact, at a time when most planetary geologists expected an equilibrium population to have a uniform slope  $b = 2$ , convinced many that the lunar highlands are *not* in equilibrium and that the present observed population is essentially equal to the production population. This result is difficult to reconcile with the evidence for very high

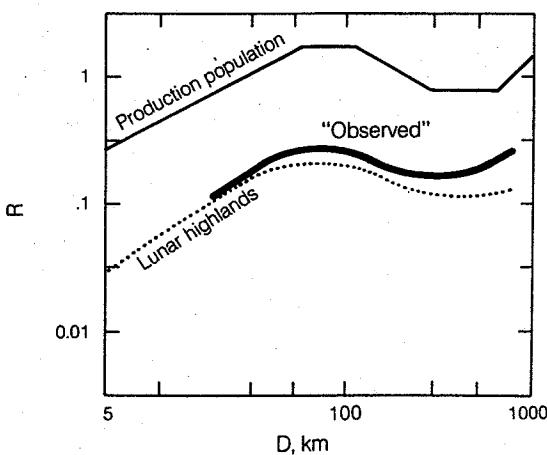
cratering rates during the period of late heavy bombardment before  $3.8 \times 10^9$  years ago, a time when cratering was so extensive that it reset the radiometric clocks of every ancient lunar rock so far returned to earth.

What the debaters failed to realize, of course, is that the concept of a constant equilibrium crater population is faulty when the production population slope  $b$  is less than 2. Figure 10.9 is the result of a simulation by Chapman using his Monte Carlo model. The observed population predicted by the model is in good agreement with the lunar highlands data on the  $R$  plot, even though the production curve is a factor 10 higher. In this simulation, nearly ten times as many craters of all diameters have been obliterated as remain on the surface, yet the observed crater population still faithfully reflects the production population. In spite of the large total number of impacts, the production population is still observable at the sites of large impacts that have erased the previous record and where only subsequent smaller impacts have been preserved. This selective preservation of smaller recent impacts also explains why the  $b > 2$  population of small impacts ( $D < 100$  m) can still be observed on the lunar surface.

### 10.3 REGOLITHS

The term *regolith*, originally denoting the fragmental layer of soil on the earth's surface, had fallen out of common geological usage by 1967 when E. M. Shoemaker revived it and applied it to the thin layer of fragmental material covering the moon's surface. In this new application, a regolith is the entire thickness of fragmental material overlying more competent rock on airless bodies. It is now known to be the product of a long period of exposure to meteoritic bombardment.

A great deal of work has been done on the nature and evolution of the lunar regolith because the Apollo missions returned many samples of it. Physically, the lunar regolith is a fine, poorly sorted powder dominated by particles less than about  $100 \mu\text{m}$  in diameter, although it contains occasional large rocks. Beginning with very porous upper layers (density less than  $1400 \text{ kg/m}^3$  in the upper centimeter), its density increases rapidly with depth because of compaction, reaching about  $2000 \text{ kg/m}^3$  at 1-m depth. Drive tube specimens have shown that the regolith is composed of many horizontal layers that do not, however, extend far laterally. These layers are presumably the ejecta blankets of small impacts. The principal constituents of the lunar regolith are rock fragments, glass, and agglutinates. This last component is characteristic of lunar soil and is an irregular accretion of rock and mineral fragments welded together by a coating of glass. Less than 1 percent of the regolith can be identified as con-



**Fig. 10.9** The results of a Monte Carlo computer calculation of the evolution of a crater population that simulates the heavily cratered lunar highlands. It is clear from this calculation that the "observed" population (that is, the population that is computed from the evolution of the production population) faithfully reflects the production population even though nearly 10 times as many craters have been obliterated in each diameter interval as exist on the surface. This final distribution is a good match to the actual lunar highlands crater population. After Chapman and McKinnon (1986).

tamination from meteorites. The composition of the regolith at each Apollo landing site reflects the composition of the underlying rock unit, suggesting that lateral transport by meteoritic impact is not very efficient. This is also supported by the relatively sharp contacts between the mare and highland units, which four billion years of meteoritic bombardment have not yet wiped out.

The color and albedo of the lunar soil is controlled by its glass content and composition. "Immature" soils are typically bright, coarse grained, and deficient in glass. As soils are bombarded by small meteorites they darken and become spectrally redder, their grain size decreases, and their glass content rises. Eventually, the soil reaches "maturity" when comminution of fragments is balanced by the formation of agglutinates. The darkening and reddening of the soil is due to the formation of dark glass that contains tiny fragments of opaque minerals and droplets of metallic iron. The lunar mare and highlands are different in brightness because of their different compositions: highlands rocks contain smaller quantities of opaque minerals and so the soils there tend to be brighter. Photometric studies show that the surface of Mercury has the low albedo and very red spectrum of mature regolith.

Although we have obtained *in situ* samples only from the lunar regolith, several types of brecciated meteorites appear to be derived from asteroidal regoliths. These "soil" samples are very immature by lunar standards, lacking much glass. Whether the low glass content is due to low gravity on the asteroidal parent bodies, which permits the escape of highly shocked crater ejecta, or whether it is due to the low (ca. 5 km/second) mean impact velocity in the asteroid belt, which is too low to induce melting, is currently under debate.

Individual mineral grains in both the lunar regolith and brecciated meteorites have been subjected to primary cosmic ray irradiation and implantation of solar wind gases, showing that even the most deeply buried components spent some time at or near the surface. Many rock fragments and mineral grains are pitted by micrometeoroid impacts, which also requires surface exposure.

Much of the work on the lunar regolith is of a petrologic or geochemical orientation and is thus outside the scope of this book. In the following section I will briefly treat those aspects of regolith evolution that most directly relate to impact processes.

### 10.3.1 Growth of the regolith

The earliest theoretical studies of the growth of the regolith by E. M. Shoemaker et al. (1969) and all subsequent studies are based on the evolution of crater populations with slopes  $b > 2$ . Although  $b$  is generally less than 2 for large craters on the lunar mare, it typically ranges between 2.9 and 3.4

for craters smaller than 4 km in diameter (Figs. 10.2 and 10.7), so the assumption that  $b > 2$  is valid for the post-mare lunar regolith and we shall consider only this case.

There has been much discussion, however, of a *megaregolith* (a term apparently coined by Hartmann, 1973) that is supposed to have developed on the surface of the moon and other planets during the era of heavy bombardment. This megaregolith is sometimes supposed to be similar to the post-mare regolith except much deeper. However, as shown in the last section, the crater population in the lunar highlands still reflects the production population and thus appears to have evolved from an initial distribution with  $b < 2$ . Since the evolution of a crater population with  $b < 2$  is very different from that with  $b > 2$ , the megaregolith may be fundamentally different from the regolith discussed below. In particular, since  $b > 2$  distributions are dominated by large impacts, the profound mixing of the surface layers seen in the post-mare regolith may not occur in the deeply brecciated rock of the megaregolith. Little work on this type of regolith has yet been done.

As shown in Figure 10.5, at any given time a  $b > 2$  crater population is in equilibrium up to some maximum crater diameter  $D_{eq}(t)$ , where  $D_{eq}(t)$  increases as  $(\text{time})^{1/(b-2)}$  for a constant rate of meteoritic bombardment. The crater population is in equilibrium with the production rate at diameters  $D < D_{eq}$  and at the smaller diameters many more craters have been obliterated than presently can be counted on the surface. Shoemaker et al. (1969) realized that an initially craterless surface subjected to this type of bombardment would rapidly develop a layer of debris with a wide variation of depths. They supposed that the maximum depth likely to occur is some fraction of the depth of the largest crater that has reached equilibrium. Although the first crater of this diameter in a given region might leave an open cavity, the ejecta blankets of subsequent craters would eventually fill it, leaving a layer of fragmental debris whose thickness is locally equal to the crater depth. This maximum likely thickness  $h_{eq}$  is taken to be

$$h_{eq} = D_{eq}/4 \quad (10.3.1)$$

The equilibrium crater diameter may be determined either directly, if it can be observed as a kink in the crater population curve (as in Figure 10.7), or indirectly, by equating the cumulative number of craters in the production population,  $N_{cum} = cD^{-b}$  to the number of craters in the equilibrium population,  $N_{eq} = c_{eq}D^{-2}$ . In this case

$$D_{eq} = (c/c_{eq})^{1/(b-2)} \quad (10.3.2)$$

where  $c_{eq}$  is dimensionless and independent of diameter, although it may depend somewhat on  $b$ . If it is supposed that equilibrium takes place at a crater density of about 4 percent of geometric sat-

## IMPACT CRATERING

uration (Eq. 10.2.7), then  $c_{eq} = 0.056$ . Conversely, the coefficient  $c$  describing the production population may be expressed in terms of  $D_{eq}$  and  $c_{eq}$ :

$$c = c_{eq} D_{eq}^{b-2} \quad (10.3.3)$$

I will use this representation of  $c$  in the rest of this section because it leads to more compact formulas without fractional powers of units.

In addition to the maximum likely regolith thickness, Shoemaker and his colleagues introduced the concept of a minimum regolith thickness. This is perhaps the least well founded concept of the model and does not agree well with the data. Unfortunately, it is also an integral part of the model. Definition of the minimum thickness begins with a straightforward integral: the fraction of the total area covered by craters with diameters between  $D$  and  $D_{eq}$ ,  $f_c(D, D_{eq})$ , given by

$$\begin{aligned} f_c(D, D_{eq}) &= -\frac{\pi}{4} \int_D^{D_{eq}} D^2 \frac{dN_{cum}}{dD} dD \\ &= \frac{\pi b c_{eq}}{4(b-2)} [(D_{eq}/D)^{b-2} - 1] \end{aligned} \quad (10.3.4)$$

where the minus sign in front of the integral is a consequence of the negative slope of the cumulative number distribution:  $dN = -(dN_{cum}/dD) dD$ .

When  $f_c(D, D_{eq})$  exceeds some fixed fraction  $f_{min}$  at diameter  $D_{min}$ , the bottoms of all craters with this diameter or less are presumed to be connected and the regolith is at least as thick as this crater's depth,  $h_{min} = D_{min}/4$ . Shoemaker et al. argued that this occurs when the target area is covered twice over by craters between  $D_{min}$  and  $D_{eq}$ , so that  $f_{min} = 2$ . Solving Equation 10.3.4 for  $f_c(D_{min}, D_{eq}) = f_{min} = 2$ , they obtained

$$h_{min} = h_{eq} \left[ \frac{4(b-2)f_{min}}{\pi b c_{eq}} + 1 \right]^{-1/(b-2)} \quad (10.3.5)$$

The probability  $P(h)$  of finding a patch of regolith of depth  $h$ , where  $h$  must lie between  $h_{min}$  and  $h_{eq}$ , is given by the ratio of the fractional area covered by craters between diameters  $D = 4h$  and  $D_{eq}$  to  $f_{min}$ :

$$P(h) = f_c(4h, D_{eq})/f_{min} \quad (10.3.6)$$

after some algebraic manipulation  $P(h)$  can be written in the convenient form

$$P(h) = \left( \frac{(h_{eq}/h)^{b-2} - 1}{(h_{eq}/h_{min})^{b-2} - 1} \right) h_{min} < h < h_{eq} \quad (10.3.7)$$

The median regolith depth  $\langle h \rangle$  occurs when  $P(\langle h \rangle) = 0.5$ . Solving the above equation it is easy to show that

$$\langle h \rangle = 2h_{eq}[(h_{eq}/h_{min})^{b-2} + 1]^{-1/(b-2)} \quad (10.3.8)$$

Table 10.1 and Figure 10.10 compare the predictions of this model to the observed regolith thickness distributions for four areas on the moon. These thicknesses were obtained from the morphology of small craters by the technique described in Section 5.6.2. Table 10.1 shows that the model tends to overestimate the regolith thickness by about 40 percent. Although some adjustments could be made to bring theory and observation into better agreement ( $f_{min} \approx 3$  would improve the fit greatly), such adjustments are probably unjustified given the other assumptions of the theory. On this level, 40 percent agreement is quite good. Figure 10.10 compares the predicted and observed probability distributions for these areas. Although Eq. 10.3.7 is a good fit at the larger depths, it fits poorly at shallow depths. This is mainly a consequence of the artificial choice of a minimum depth, although it may also be due to neglect of the previously formed regolith in blanketing the surface (Gault, 1970).

Another important result of this model is that it predicts that the regolith thickness may increase nonlinearly with time even if the cratering rate is constant. Both  $h_{eq}$  and  $h_{min}$  are related to  $D_{eq}$  by constant factors for constant  $b$ , so the median depth  $\langle h \rangle$  is also proportional to  $D_{eq}$ . Since  $D_{eq}$  increases as  $(\text{time})^{1/(b-2)}$  for constant cratering rate, so do all these measures of the regolith thickness. Only if  $b = 3$  does the regolith accumulate at a constant rate. Data on the crater population, regolith thickness, and ages of the underlying mare at the Apollo 11 landing site indicates  $b \approx 2.9$  so that the lunar regolith there accumulated at a nearly constant rate of about  $1.2 \text{ m}/10^9 \text{ years}$ .

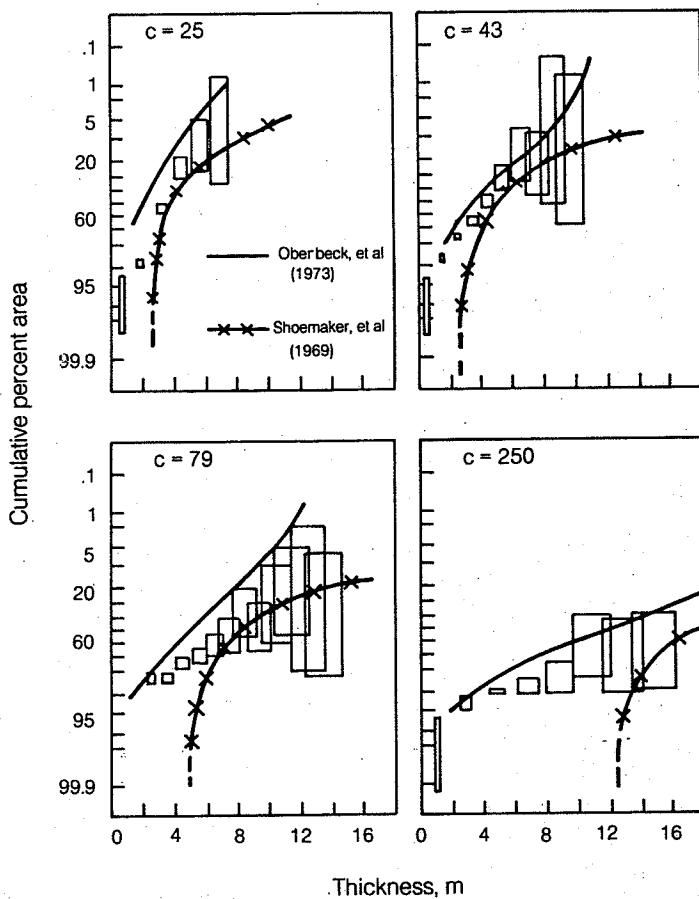
TABLE 10.1 Predicted vs. observed regolith thickness

Area*	$D_{eq}\dagger$	Shoemaker et al. (1969)			$\langle h \rangle$	$\langle h \rangle$	Observed‡ $\langle h \rangle$
		$h_{eq}$	$h_{min}$	$\langle h \rangle$			
LO III P-11	80	20	2.3	4.5	3.3	3.3	
LO III P-13b	120	29	3.4	6.6	4.9	4.6	
LO II P-7b	180	45	5.3	10.0	7.1	7.5	
LO V 24	410	100	12.0	23.0	15.0	16.0	

\*Areas are designated by Lunar Orbiter mission and frame. For more detail see Oberbeck and Quaide (1968).

†Computed from Equation 10.3.2 using  $c_{eq} = 0.056$ ,  $b = 3.4$  for all sites and  $c = 25, 43, 79$ , and  $250$  respectively for the four sites, where  $c$  is in units of  $\text{m}^{1.4}$ .

‡From Oberbeck and Quaide (1968).



**Fig. 10.10** Comparison of the statistical variation of measured lunar regolith thicknesses with the theoretical models of Shoemaker et al. (1969) (data from Table 10.1 and curves from Eq. 10.3.7) and the Monte Carlo computer calculation of Oberbeck et al. (1973). The four different areas are characterized by different production populations (cumulative crater density parameter  $c$  is in units of meters<sup>1/4</sup>). Although the Shoemaker et al. (1969) model is a poor fit for the thinner regions of the regolith, it represents the probability of finding deeper regions well. Modified after Oberbeck et al. (1973).

On the whole, the regolith evolution model of Shoemaker et al. (1969) gives an adequate first-order description of the growth of the regolith. It has the great advantage of being analytic, so that the equations may be easily applied to many different situations, and it relies on the observed crater population for its input parameters. Although it might be improved by a better statistical treatment of the smaller regolith thicknesses, no one has yet published a revised model.

A second regolith evolution model was developed by V. R. Oberbeck et al. (1973). This is a numerical model that uses Monte Carlo techniques, so it is not easily generalized to areas that were not specifically treated by the authors. Oberbeck and his colleagues emphasize the role of ejecta blankets in forming the regolith, in contrast to the emphasis by Shoemaker and his colleagues on crater depth. In their simulation Oberbeck and his colleagues assume that the ejecta from a crater is spread out to a radius of seven times the crater

radius. Although various ejecta distributions were studied, they preferred one in which the ejecta thickness falls linearly from a maximum at the crater rim to zero at the assumed edge of the blanket. The use of a linear instead of a power-law distribution is an attempt to account for downslope transport of the originally sharp crater rim by impacts too small to include in the computation. The model also takes into account the effect of the strength transition at the base of the regolith on crater morphology, thus including the blanketing of bedrock by regolith. The major input parameter of this model is the observed crater population on the regolith.

The results of this model suggest that more than half of the regolith is built up of the overlapping ejecta blankets of craters larger than the equilibrium diameter. The Monte Carlo model also showed that the regolith depth does not grow linearly with time, but that its depth increases roughly as  $(\text{time})^{0.7}$ . This apparent disagreement

with Shoemaker et al. (1969) is due to use of a production population slope  $b = 3.4$  by Oberbeck and his colleagues. If this slope is substituted into the formula of Shoemaker and his colleagues, then exactly the same dependence of depth on time is obtained. Table 10.1 and Figure 10.10 show the Monte Carlo model predictions for mean regolith thickness and thickness distribution. The model does a somewhat better job than the analytic model on mean thickness and gives a much better fit to the observed thickness distributions for small thicknesses. The two models, although they give different results, fit the data equally well for large thicknesses.

### 10.3.2 Mixing of the regolith

Mixing of the regolith by meteoritic impact is closely related to its growth, except that the regolith can only grow to a given depth once, whereas it may later be overturned to this depth many times. Since small craters are more abundant than large ones, it is obvious that the regolith will be overturned more frequently near the surface than at depth. This process has been referred to by the apt term *gardening*, meaning the repeated stirring and overturn of the regolith once it has formed. Rare gas and particle track studies have shown that large fractions of the moon's regolith have spent some time at or near the surface before being buried by the ejecta blanket of some nearby crater. Impacts into the porous regolith may also push individual mineral grains that originally lay more or less directly under the impact deeper into the regolith.

Many models have been constructed to describe mixing of the regolith. Most of these models focus on some special aspect of regolith evolution to explain data obtained from the Apollo or Luna sample returns. Two general models will be discussed here. The first is analytic and is closely related to the work of Shoemaker et al. (1969) described in the last section. The second is statistical and was devised by D. E. Gault et al. (1974). The first model has not, in my opinion, received the attention it deserves, whereas the second, although it has had a great deal of influence on subsequent work, contains an error that led to incorrect conclusions about overturn rates.

E. M. Shoemaker et al. (1970) extended the analytic model described in the last section to obtain estimates of the regolith turnover times to various depths. Although "turnover" can be defined in different ways, Shoemaker and his colleagues (a different and larger cast than in 1969—this is a symptom of "big science" in the Apollo era) chose a prescription that equates "turnover" with "disturb." Thus, if it is presumed that a crater of diameter  $D$  disturbs regolith material to a depth  $d = D/4$ , then the length of time, on average, that passes between subsequent "turnovers" is roughly

equal to the time that must elapse before the fractional area covered by craters of diameter  $D = 4d$  or larger (up to  $D_{eq}$ ) equals one. Since the age of the regolith itself is roughly the time for craters equal to or larger than  $D_{eq}$  to cover the surface once, the turnover time can be expressed in terms of the age of the surface  $t_{surf}$

$$t_{ovr} = t_{surf}/f_c(4d, D_{eq}) \quad (10.3.9)$$

A change in the definition of "turnover" requires only a change in the first argument of  $f_c$ . For example, if "turnover" is supposed to mean excavation and redeposition on the surface, then the excavation depth  $h_{exc} = D/10$  (Eq. 5.5.4) should be used in Equation 10.3.9. The discerning reader will note that there is a problem with Equation 10.3.9 because the overturn time to depth  $D_{eq}$  is apparently infinity, not  $t_{surf}$ . In actuality the upper limit of integration in  $f_c$  should be infinity instead of  $D_{eq}$ , which would cure the problem. However, this flaw is not serious so long as  $t_{ovr} \ll t_{surf}$ , and I will follow the original authors in leaving the upper limit  $D_{eq}$ . The interested reader may be able to make other improvements in this scheme.

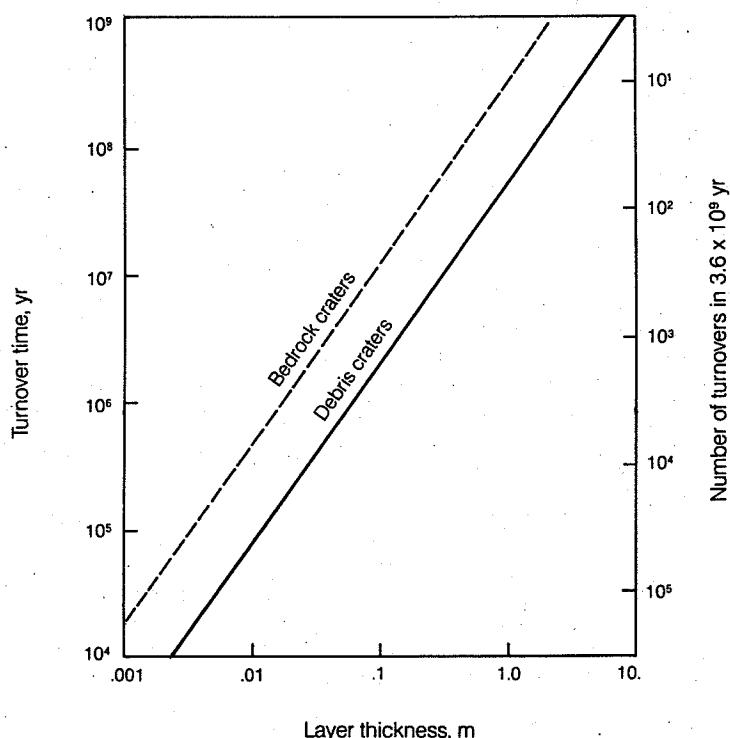
Using Equation 10.3.4, the overturn time  $t_{ovr}$  is related to parameters describing the crater population on the surface by

$$t_{ovr} \approx \frac{4(b-2)}{\pi b c_{eq}} (d/h_{eq})^{b-2} t_{surf} \quad (10.3.10)$$

where the approximation  $d \ll h_{eq}$  has been used in simplifying this equation. Figure 10.11 illustrates this equation's predictions for the Apollo 11 Tranquillity Base site. It is notable that the upper millimeter has suffered about 30 times as many overturns at this site as the upper centimeter, which itself has been overturned about 30 times more than the upper 10 cm.

The second mixing model was constructed by D. E. Gault et al. (1974). Unlike the other models I have discussed, this model does not use the observed crater population to provide input parameters but instead uses the postulated projectile population and scaling laws to obtain a predicted crater production population. Once the production population has been computed, this model is rather similar to the model of Shoemaker et al. (1969) discussed above, except that Gault et al. presume that "turnover" means "excavation" and use an excavation depth  $h_{exc} = D/8$ . The Gault et al. (1974) model also handles the statistics of cratering more carefully.

The results of this model, particularly the plot of number of times of turnover as a function of depth and age, have been widely cited and have greatly influenced the general perception of regolith processes. However, one aspect of these results has been very puzzling, and that is the prediction that the overturn time to depth 10 cm is about  $10^9$  years. Gault et al. (1974) suggested that



**Fig. 10.11** Regolith overturn time as a function of depth. The overturn time is computed from the theory of Shoemaker et al. (1970) and Equation 10.3.10 using parameters measured at the Apollo 11 Tranquillity base landing site, where the mean regolith thickness is 4.1 m and the underlying mare is  $3.6 \times 10^9$  years old. The right-hand scale shows the number of turnovers at each depth since the formation of the regolith. The dashed curve is computed directly from the theory. The solid curve is supposed to be a better representation of reality, stemming from the suggestion of Shoemaker et al. that craters formed in loose regolith are approximately a factor six times larger than craters in bedrock, thus hastening the rate of overturn in loose material by comparison with the rate of formation from bedrock.

the high meteoritic flux during the waning stages of heavy bombardment might have allowed deeper mixing in the past, although the present surface is only overturned at the calculated rate. This proposal has found little support from studies of exposure of the regolith to cosmic rays, which indicate that rocks reside in the upper 10 cm only about  $5 \times 10^8$  years rather than the  $3 \times 10^9$  years predicted by the Gault et al. model. The problem in this model appears to be with the scaling relation relating crater diameter and the projectile diameter. Gault et al. used an energy scaling law that produces a predicted production population with a slope  $b = 4.14$ —far steeper than any observed production population on the lunar surface. A steeper production population means more small impacts for a given number of large impacts, and it causes the overturn time to rise more rapidly with increasing depth. If the scaling relation of Gault and his colleagues is replaced with the modern Pi-group scaling parameters for loose sand (Sec. 7.7 and Table 7.1), then the predicted production population slope becomes  $b = 2.97$ , in remarkably good agreement with the value  $b = 2.93$  for Tranquillity Base of

Shoemaker and his colleagues. With this correction the model of Gault and his colleagues makes predictions more in line with the observed residence times for rocks in the lunar regolith.

### 10.3.3 Horizontal Transport

Both the Shoemaker et al. (1970) and Gault et al. (1974) models agree in predicting that horizontal transport on the lunar surface is a very slow process. Some early studies used data on ejection velocity from craters formed in strong substrates (strength regime; see Sec. 7.7) to predict that even very small craters would throw ejecta distances of several kilometers, resulting in horizontal transport rates many orders of magnitude larger than vertical mixing rates. However, the recent advances in understanding ejecta distributions of the gravity regime craters that form in loose regolith material (Sec. 7.12) show that ejected debris moves horizontally only a few times farther than the crater depth, so that horizontal and vertical mixing rates are roughly comparable. This does not, of course, prohibit large, rare impact events from widely dispersing small amounts of mate-

rial: indeed, small quantities of what is believed to be Tycho ejecta was found in Apollo 17 samples, and a fragment of Copernicus ejecta is thought to have been found at the Apollo 15 landing site. However, most material in the regolith merely executes a random drunkard's walk with hops of a few centimeters or at most meters and thus never gets far from the site where it was initially torn from the bedrock. This lack of lateral mixing has preserved the relatively sharp contacts between mare and highland rock types in spite of the formation of a regolith on both surfaces.

An analytic model describing the rate of horizontal transport was constructed by Arvidson et al. (1975). These authors began with the observation that the ejecta blankets of regolith craters scale self-similarly (see Secs. 6.2 on crater rims and 7.12 on ejecta scaling) and that the blanket thickness  $\delta$  is a function of distance from the crater center  $r$  and crater diameter  $D$ . Rewriting Equation 6.3.2 slightly, and reducing the coefficient 40 percent to account for rim uplift, we find the thickness of primary ejecta deposited around a crater of diameter  $D$  to be roughly

$$\delta(r, D) = 0.10(D/2)^{0.74}(2r/D)^{-3} \quad (10.3.11)$$

Note that the rims of very small regolith craters (less than about 1 m in diameter) do not seem to obey this relation, so all of the following discussion will be restricted to craters larger than 1 m. Note further that Eq. 10.3.11 ignores any effect

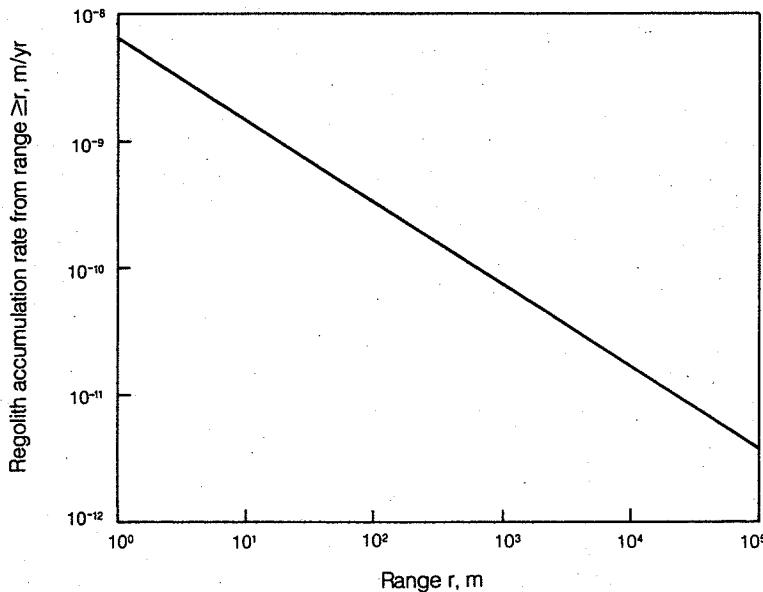
surface gravity  $g$  might have on rim thickness. Gault and Wedekind (1979) report that  $\delta$  varies approximately as  $g^{1/6}$ , so lunar rim thicknesses would be about 75 percent of their terrestrial counterparts.

Now consider some point on the regolith surface. The number of craters with diameters between  $D$  and  $D + dD$  impacting at a distance between  $r$  and  $r + dr$  from this point is the number per unit area in this interval,  $dN = -(dN_{\text{cum}}/dD)dD$ , times the area  $2\pi r dr$ . Since the thickness of ejecta deposited by each such crater is  $\delta(r, D)$ , the total thickness  $dT(r)$  deposited by all craters impacting between  $r$  and  $r + dr$  is

$$dT(r) = -2\pi r dr \int_0^{2r} \delta(r, D) \frac{dN_{\text{cum}}}{dD} dD \quad (10.3.12)$$

where the integral is cut off at a diameter  $D$  equal to the range  $2r$ , since a larger crater excavates the point instead of depositing ejecta upon it. The cumulative thickness of ejecta  $T(\geq r)$  that is deposited from a distance equal to or greater than  $r$  is simply

$$\begin{aligned} T(\geq r) &= \int_r^{\infty} dT(r') \\ &= -2\pi \int_r^{\infty} r' \left[ \int_0^{2r'} \delta(r', D) \frac{dN_{\text{cum}}}{dD} dD \right] dr' \end{aligned} \quad (10.3.13)$$



**Fig. 10.12** Rate of accumulation of regolith as a function of the distance from which the material originates. This figure is constructed from the horizontal transport theory of Arvidson et al. (1975) using slightly different parameters than those of the original authors (Eq. 10.3.14). The results show that more than 99 percent of the material at a given site originates from distances of less than 1 km. This model neglects a small but significant contribution from the secondary ejecta and rays of large distant impacts.

Using Equation 10.3.11 and cumulative crater density (Equation 10.2.2) the integral of Equation 10.3.13 can be computed for arbitrary  $c$  and  $b$ :

$$T(\geq r) = \frac{\pi b c}{5(3.74 - b)(b - 2.74)2^b} r^{2.74-b} \quad (10.3.14)$$

Note that to be valid  $2.74 < b < 3.74$ . Slopes outside this range cause divergences at one or another of the integration limits.

The rate of regolith accumulation from distances equal to or greater than  $r$  is obtained from this same equation if  $c$  is replaced by its time derivative. Data from the lunar landing sites suggest that  $dc/dt \approx 7.6 \times 10^{-9} \text{ m}^{1.4}/\text{yr}$  and  $b \approx 3.4$ , resulting in

$$\frac{dT}{dt} (\geq r) = 6.8 \times 10^{-9} r^{-0.66} \quad (10.3.15)$$

where  $dT(\geq r)/dt$  is in units of meters per year and  $r$  is in units of meters. The total accumulation rate at some point on the regolith from impacts one or more meters away is 6.8 mm/Myr.

Figure 10.12 plots the rate of accumulation (Eq. 10.3.15) as a function of range for the parameters quoted. It is clear from this plot that most of the material at a given site (more than 99 percent) originates within 1 km. This model, of course, depends heavily on the accuracy of the ejecta thickness distribution (Eq. 10.3.11). Although this distribution is adequate for most of the ejecta, secondary craters and rays, discussed in Sections 6.4 and 6.5, are dispersed much more widely so that a small amount of material at each site may have been derived from distant craters by the processes that produce these ejecta types.

## FURTHER READING

### General References

The two general areas of crater statistics and regolith evolution have been extensively treated in many publications. One of the best and most complete recent reviews of work on crater statistics and age dating is in an unlikely sounding place:

Basaltic Volcanism Study Project, *Basaltic Volcanism on the Terrestrial Planets*, Pergamon Press, New York, Chapter 8 (1981).

An excellent and comprehensive review of regolith evolution is given by

Y. Langevin and J. R. Arnold, The evolution of the lunar regolith, *Ann. Rev. Earth Planet. Sci.* 5: 449-489 (1977).

### Special Topics

The standards for collection and presentation of

crater statistical data were established by a NASA panel composed of an international group of experts. Their recommendations are published in

*Standard Techniques for Presentation and Analysis of Crater Size-Frequency Data*, NASA Technical Memorandum 79730 (1978).

The use of craters to date planetary surfaces is particularly well explained by

G. E. McGill, Craters as "fossils": The remote dating of planetary surface materials, *Geol. Soc. Amer. Bull.* 88: 1102-1110 (1977).

One of the classic papers on this subject still deserves to be read:

E. M. Shoemaker, R. J. Hackman, and R. E. Eggleton, Interplanetary correlation of geologic time, *Adv. in Astronaut. Sci.* 8: 70-89 (1963).

One of the most complete recent summaries of the primary meteoroid flux for objects less than about 10 cm in diameter is found in the first part of

E. Grün, H. A. Zook, H. Fechtig, and R. H. Giese, Collisional balance of the Meteoritic Complex, *Icarus* 62: 244-272 (1985).

The flux of larger meteoroids is treated in Gault et al. (1974), referred to below in a different context. The crater production rate on the earth is carefully examined by

R. A. F. Grieve and M. R. Dence, The terrestrial cratering record II. The crater production rate, *Icarus* 38: 230-242 (1979).

Research on the evolution of crater populations has a complicated history with many papers falling just short of clarifying the major issues. The best paper on the evolution of populations with  $b > 2$  is, without a doubt,

D. E. Gault, Saturation and equilibrium conditions for impact cratering on the lunar surface: Criteria and implications, *Radio Science* 5: 273-291 (1970).

An excellent analytic approach to the evolution of crater populations with  $b > 2$  was made by

L. A. Soderblom, A model for small-impact erosion applied to the lunar surface, *J. Geophys. Res.* 75: 2655-2661 (1970).

This model demonstrated both the evolution of an equilibrium population with a slope  $b = 2$  and the dependence of the equilibrium density on  $b$ .

The recent work clarifying the evolution of crater populations with  $b < 2$  has, so far, only appeared as part of a review!

C. R. Chapman and W. B. McKinnon, Cratering of planetary satellites, in J. A. Burns and M. S. Matthews (Eds.), *Satellites*, University of Arizona Press, Tucson, pp. 492-580 (1986).

A fine review of the petrology and physical characteristics of lunar soil is given by

## IMPACT CRATERING

G. Heikin, Petrology of lunar soils, *Rev. Geophys. Space Phys.* 13: 567-587 (1975).

The first detailed analysis of the growth of the lunar regolith appears as part of a much longer analysis of the Surveyor 7 data:

E. M. Shoemaker, R. M. Batson, H. E. Holt, E. C. Morris, J. J. Rennilson, and E. A. Whittaker, Observations of the lunar regolith and the earth from the television camera on Surveyor 7, *J. Geophys. Res.* 74: 6081-6119 (1969).

A Monte Carlo analysis of regolith growth that is still cited as the classic work in the field appeared somewhat later:

V. R. Oberbeck, W. L. Quaide, M. Mahan, and J. Paulson, Monte Carlo calculations of lunar regolith thickness distributions, *Icarus* 19: 87-107 (1973).

The analytic model of regolith mixing described in the text is developed in

E. M. Shoemaker, M. H. Hait, G. A. Swann,

D. L. Schleicher, G. G. Schaber, R. L. Sutton, D. H. Dahlem, E. N. Goddard, and A. C. Waters, Origin of the lunar regolith at Tranquillity Base, *Proc. Apollo 11 Lunar Sci. Conf.*, 3: 2399-2412 (1970).

One of the most influential papers on the mixing of the regolith also contains one of the best current estimates of the projectile population in the earth-moon system:

D. E. Gault, F. Horz, D. E. Brownlee, and J. B. Hartung, Mixing of the lunar regolith, *Proc. 5th Lunar Sci. Conf. (Supplement Geochim. Cosmochim. Acta)*, 5: 2365-2386 (1974).

Lateral transport of the lunar regolith is well treated by

R. Arvidson, R. J. Drozd, C. M. Hohenberg, C. J. Morgan, and G. Poupeau, Horizontal transport of the regolith, modification of features, and erosion rates on the lunar surface, *The Moon* 13: 67-79 (1975).

## XI. ATMOSPHERIC INTERACTIONS

### 11.1 INTRODUCTION

Most studies of impact cratering have focused on lunar, Martian, or Mercurian craters and so tended to neglect the role of atmospheric gases in the cratering process. However, the discovery of increasing numbers of terrestrial impact craters and the recognition of craters on Venus makes a consideration of atmospheric interactions imperative in any extensive treatment of impact cratering.

The effect of an atmosphere on the formation of impact craters can be seen as twofold. The first, which is by far the best studied, is related to the projectile's flight through the atmosphere. For sufficiently small projectiles this is also the end of the story, because the projectile may be completely vaporized upon atmospheric entry or slowed so greatly that it merely falls to the surface at terminal velocity. Larger projectiles may be fragmented and distorted by their passage through the atmosphere, producing an impact crater morphology not seen on airless bodies. The second aspect of the atmospheric interaction occurs after a large enough projectile has struck the surface and the vapor and ejecta thrown from the crater meet the ambient atmosphere. This type of interaction has not yet received a great deal of study, although the recent recognition that large impacts might have been responsible for major biological extinctions on earth has spurred research in this area.

### 11.2 PENETRATION OF THE ATMOSPHERE

The minimum encounter velocity between a planet and a meteoroid is the planet's escape velocity. The maximum is a combination of its escape velocity, heliocentric orbital velocity, and the velocity of an object just barely bound to the sun at the planet's orbital position. For the earth the minimum velocity is 11.2 km/second and the maximum is 72.8 km/second. When a meteoroid collides with the earth or some other planet with a substantial atmosphere it thus encounters atmospheric gases at high speed. Whether it will retain its high velocity after traversing the atmosphere depends upon its mass or diameter and the thickness of the atmosphere.

A rough approximation can be used to gauge the ability of a given meteoroid to penetrate the

atmosphere. As the meteoroid traverses the atmosphere it deflects a volume of gas roughly equal to the meteoroid's cross-sectional area times the length of its trajectory. If the mass of gas in this volume is more than ten times the mass of the meteoroid, its velocity is reduced to less than 10 percent of its initial value by momentum conservation and can be considered "stopped." Since the integrated mass of atmospheric gas along a trajectory sloping at angle  $\theta$  is  $P_{\text{surf}}/(g_{\text{surf}} \sin \theta)$  per unit area, where  $P_{\text{surf}}$  is the surface atmospheric pressure and  $g_{\text{surf}}$  is the surface acceleration of gravity, for a spherical projectile of diameter  $L$  this criterion is

$$L \geq 0.15 \frac{P_{\text{surf}}}{\rho_p g_{\text{surf}} \sin \theta} \quad (11.2.1)$$

where  $\rho_p$  is the meteoroid's density. Table 11.1 gives the minimum diameter projectile that can penetrate an atmosphere at vertical incidence ( $\theta = 90^\circ$ ) for Venus, earth, Mars, and Titan, computed from Equation 11.2.1 using densities appropriate for ice, rock, and iron meteoroids. It may seem that a minimum diameter of 60 m for stony meteoroids striking the earth is too large, since recovered meteorites are much smaller. However, the penetration criterion is concerned with the minimum-diameter meteoroid that can traverse the atmosphere and retain sufficient velocity to produce a *hypervelocity* impact crater—and the numbers in the table are roughly consistent with observation of terrestrial craters of this type. Meteorite fragments are almost all recovered from meteoroids that have been "stopped" by the atmosphere and that fall subsonically to the earth. The "craters" produced by such falling objects are created by a different set of processes than those produced by high-velocity impacts.

#### 11.2.1 Flight through the atmosphere

The detailed process of meteoroid entry into an atmosphere is quite well understood. The meteoroid first encounters the very tenuous upper atmosphere, where the mean free path of the molecules exceeds the meteoroid's diameter in the "molecular flow" regime. Molecules of the atmosphere strike the meteoroid independently and either stick or rebound, producing a small amount of retardation at the same time as they sputter

## IMPACT CRATERING

TABLE 11.1 *Minimum diameters of meteoroids that can penetrate an atmosphere\**

Planet	$g_{\text{surf}}$ (m/sec <sup>2</sup> )	$P_{\text{surf}}$ (MPa)	$L_{\text{min}}(\text{ice})$ (m)	$L_{\text{min}}(\text{stony})$ (m)	$L_{\text{min}}(\text{iron})$ (m)
Venus	8.6	9.0	160.000	58.000	20.000
Earth	9.8	0.1	1.500	.600	.200
Mars	3.7	0.0006	.024	.009	.003
Titan	1.44	0.16	17.000	6.000	2.000

\*Computed from Equation 11.2.1 in text

molecules off the surface of the meteoroid and heat it. This flow regime is of negligible importance for large meteoroids, but tiny micrometeoroids (less than 10 to 100  $\mu\text{m}$  in diameter, depending upon angle and velocity of incidence) may be slowed sufficiently gradually that they can radiate away the acquired heat and come to rest nearly unaltered in the earth's upper atmosphere. The collection of such "cosmic dust" has now become a major scientific enterprise and a valuable tool in the investigation of the solar system.

Larger meteoroids are not so easily slowed and soon reach the denser layers of the atmosphere, which they traverse at Mach numbers (the ratio between meteoroid speed and the local speed of sound) often exceeding 50. A bow shock develops in front of such a meteoroid where atmospheric gases are compressed and heated in a narrow "standoff" zone between the shock front and the meteoroid's surface. The hot gases in turn radiate their heat to the surface of the meteoroid, which responds by a combination of melting and vaporization, collectively ablating away the meteoroid's surface as air flow carries this material into the wake. The pressure of the compressed gases in front of the meteoroid is approximately the stagnation pressure,  $\rho_g v^2$ , where  $\rho_g$  is the density of the ambient gas and  $v$  is the meteoroid's speed. The pressure behind the meteoroid in its wake is near zero. This differential pressure across the meteoroid's body produces a net drag force that decelerates it.

The combined forces of drag, lift (if present), and gravity in the presence of ablation result in a set of differential equations that completely describe the trajectory of a meteoroid entering a planet's atmosphere. There are five major equations and two subsidiary ones. Some of the geometric parameters are illustrated in Figure 11.1. The first equation describes the change in speed that is due to drag and gravitational acceleration  $g$  (which must be considered as a function of altitude  $Z$  in deep atmospheres)

$$\frac{dv}{dt} = -C_D \frac{\rho_g A v^2}{m} + g \sin \theta \quad (11.2.2)$$

where  $C_D$  is the drag coefficient,  $A$  is cross-sectional area, and  $m$  is the meteoroid mass. The minimum velocity that the meteoroid can attain

occurs when all of its initial momentum has been spent by drag and it falls vertically toward the surface. Under these circumstances  $dv/dt = 0$  because the gravitational acceleration is balanced by aerodynamic drag. The velocity at which this occurs is known as the terminal velocity  $v_{\text{term}}$  and Equation 11.2.2 can be used to show that  $v_{\text{term}} = (mg/C_D \rho_g A)^{1/2}$ .

The next equation describes the rate of change of mass  $m$  resulting from ablation:

$$\frac{dm}{dt} = -\frac{C_H \rho_g A v^3}{2\zeta} \left( \frac{v^2 - v_{\text{cr}}^2}{v^2} \right) \quad (11.2.3)$$

where  $\zeta$  is the heat of ablation of the meteoroid material, about  $5 \times 10^6 \text{ J/kg}$  for both stony and iron meteorites, and  $C_H$  is the heat transfer coefficient, which is somewhat uncertain but a value 0.02 seems to be a good average.  $v_{\text{cr}}$  is the critical velocity below which ablation is unimportant. It has been experimentally determined to be about 3 km/second for iron meteorites. At velocities  $v$  below  $v_{\text{cr}}$ , ablation does not occur and  $dm/dt = 0$ .

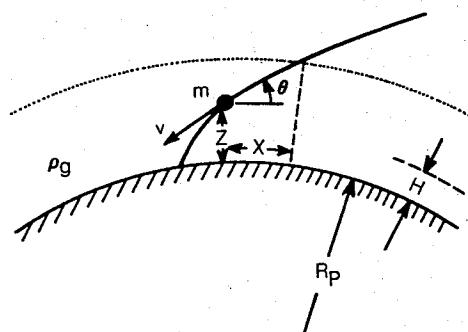


Fig. 11.1 Variables used to describe the path of a meteoroid of mass  $m$  as it descends through the atmosphere of a planet with radius  $R_p$ . The atmospheric density at any given altitude is  $\rho_g$  and its scale height is  $H$ . The meteoroid's altitude is  $Z$ , its instantaneous speed is  $v$ , and its trajectory forms an instantaneous angle  $\theta$  with the local horizontal. The meteoroid's distance downrange from an arbitrary initial position is  $X$ . In this figure the initial position is determined by the intersection between the meteoroid's trajectory (solid line) and an arbitrary shell located well outside the planet's atmosphere (dotted line).

The trajectory angle  $\theta$  is given by

$$\frac{d\theta}{dt} = \frac{mg \cos \theta - \frac{1}{2} C_L \rho_g A v^2}{mv} - \frac{v \cos \theta}{R_p + Z} \quad (11.2.4)$$

where  $C_L$  is the lift coefficient (if lift is present—the tumbling of meteoroid fragments generally prevents this term from exceeding 0.001),  $R_p$  is the radius of the planet, and  $Z$  is altitude above the surface. The second term in this equation may be neglected if the planet's surface can be approximated as flat. It is only needed for calculations with very low entry angles.

The last two equations are kinematical and merely relate speed  $v$  and angle  $\theta$  to altitude  $Z$  and downrange distance  $X$ :

$$\frac{dZ}{dt} = -v \sin \theta \quad (11.2.5)$$

and

$$\frac{dX}{dt} = \frac{v \cos \theta}{1 + Z/R_p} \quad (11.2.6)$$

Auxiliary equations give ambient gas density  $\rho_g$  as a function of altitude (although this equation is approximate, it is generally good enough for atmospheric entry computations):

$$\rho_g = \rho_0 \exp(-Z/H) \quad (11.2.7)$$

where  $H$  is the atmospheric scale height near the surface ( $H$  is roughly equal to 10 km on Venus, earth, and Mars) and  $\rho_0$  is the surface atmospheric density. Finally, the cross-sectional area  $A$  is a function of meteoroid mass, decreasing as material is ablated away. It is generally given as

$$A = S_F (m/\rho_p)^{2/3} \quad (11.2.8)$$

where the shape factor  $S_F$  is 1.2 for spheres.

Solution of these coupled differential equations is usually accomplished by a fourth-order Runge-Kutta technique and results in reasonably accurate predictions for the trajectory of a meteoroid that travels through the atmosphere without breaking up.

### 11.2.2 Shock wave interactions

As a meteoroid traverses the atmosphere it also generates a strong shock wave and at high velocities may emit a bright light. This light, which is mainly emitted by hot atmospheric gas and meteoritic vapor and melt droplets, is responsible for the phenomenon of visual meteors. The shock waves produced by visual meteors are seldom observed on the ground because most such meteors burn up at high altitude (most visual meteors are produced by meteoroids of only a few grams mass). However, the shock waves of large meteoroids that strike the ground at high velocity may themselves be quite devastating. The shock wave produced by the 1908 Tunguska meteoroid leveled about two thousand square kilometers of dense Siberian forest, snapping off meter-diameter trees like matchsticks, in spite of the fact that the meteoroid did not quite reach the ground. A trader 60 km from the impact site was knocked to the ground by the shock and his clothes charred by thermal radiation from the incandescent meteor. The mass of the Tunguska meteoroid is estimated to have been  $3.5 \times 10^9$  kg, just a little above the minimum size for a meteoroid that can penetrate the earth's atmosphere.

The energy dissipated by a meteoroid as it traverses the atmosphere is simply the integral of the drag force over the length of its trajectory. For a meteoroid so large that ablation and its velocity decrease because of drag can be neglected, the atmospheric energy deposition  $E_{atm}$  is

$$E_{atm} = (\pi/2) C_D L^2 v^2 \rho_0 H / \sin \theta \quad (11.2.9)$$

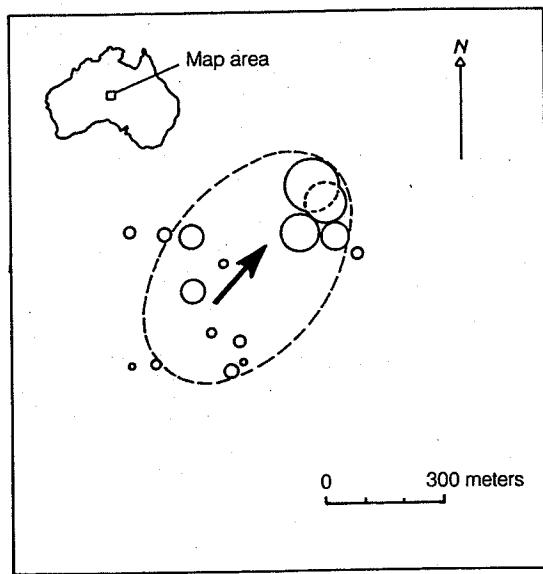
Small meteoroids, of course, deposit all of their initial kinetic energy in the atmosphere. This deposited energy is partitioned mainly into shock waves and heat; light emission accounts for only an estimated 0.01 percent of the total energy consumed.

### 11.2.3 Aerodynamic breakup

Equations 11.2.2 through 11.2.8 neglect a major feature of meteorite flight through an atmosphere because they assume that the meteoroid remains intact as a single object as it traverses the atmosphere. Observed meteors, however, exhibit sudden flareups of unusually large light emission that are attributed to fragmentation events. Small meteoroids that strike the ground seldom fall singly, but instead fall in a group or cluster known as a *crater field* or *strewn field*. Figure 11.2 illustrates one such cluster near Henbury Cattle Station in central Australia. In the great 1947 Sikhote-Alin meteorite fall in Siberia over 150 craters ranging in diameter from 26.5 m to about 0.1 m were formed by iron meteorite fragments that fell within an elliptical area 2-km long and 1-km wide.

These observations indicate that meteoroids are frequently broken up as they traverse the atmosphere. This is not surprising, as aerodynamic stresses, roughly equal to  $\rho_g v^2$ , rise from about 10 MPa for a 20 km/second meteoroid at 30 km altitude on earth to nearly 100 MPa at an altitude of 15 km. The observed crushing strengths of recovered stony meteorites (which have survived not only atmospheric passage but also impact with the earth's surface) typically range from about 6 MPa to 500 MPa. The strengths of carbonaceous chondrites may be only about 1 MPa (which is probably why carbonaceous chondrites are seldom recovered as large fragments), and typical visual meteorites break up at stresses of only 0.002 MPa.

## IMPACT CRATERING



**Fig. 11.2** Schematic map of the Henbury crater field in Australia. The dashed ellipse measures  $440 \times 640$  meters, and the arrow shows the inferred flight direction. *Modified from Passey and Melosh (1980).*

Most meteorites entering the earth's atmosphere are thus expected to break up unless they burn up first. Only the toughest iron or stony meteorites can penetrate the atmosphere at high speed without being crushed by aerodynamic stresses. On Mars, however, with its much less dense atmosphere, because aerodynamic stresses are a factor 130 times less than on the earth most meteoroids should be able to survive Martian atmospheric passage without fragmentation. On the other hand, Venus's surface atmospheric density is about five times larger than earth's, so aerodynamic stresses are five times greater. It thus is unlikely that any type of meteoroid can reach Venus's surface without breaking up first.

It is important to note that the breakup of meteoroids entering an atmosphere does not depend upon size: because aerodynamic stresses are the same for both large and small meteoroids, if they have equal crushing strengths both sizes are equally liable to break into pieces. In fact, the statistical theory of strength suggests that, on average, larger meteoroids will be somewhat weaker than small ones.

Although *all* sizes of entering meteorites will thus be broken up in the earth's atmosphere, the outcome of the breakup does depend upon the size of the meteoroid. Figure 11.3 schematically illustrates the evolution of a fragmenting meteoroid. The meteoroid first enters the atmosphere and, as it encounters progressively denser atmospheric gases, stresses rise steadily. Although particularly weak fragments may break off at high altitude, most crushing takes place in the lower

atmosphere. As the main body of the meteoroid is crushed it expands laterally, increasing its surface area and hence drag, which increases the stress on the meteoroid still further. The mass of crushed fragments continues to expand laterally until the bow shock enclosing the fragments dissolves into separate bow shocks, each enclosing a separate fragment. The individual fragments may themselves break up, or they may continue intact until they strike the ground.

When an incoming meteoroid progresses through the entire sequence of steps illustrated in Figure 11.3 before it strikes the ground, the result is a crater field of the type illustrated in Figure 11.2. In such fields the largest crater or craters is usually at one end of the ellipse. Smaller fragments fall short of the larger ones, indicating that they have been slowed more by drag and thus do not travel as far downrange. The scatter ellipse also has a finite width that indicates that the fragments have received a cross-range velocity impulse. This impulse is apparently imparted at the time of breakup and is due to a combination of the lateral expansion velocity of the original meteoroid before it broke down into separate fragments and a repulsive bow shock interaction that blasts fragments apart as the single bow shock that originally enclosed them dissolves into separate bow shocks. For a short interval, as the fragments separate, the pressure between them is larger than the stagnation pressure, forcing the pieces apart. This interaction ceases when they separate a distance comparable to their diameters. The transverse velocity thus imparted,  $v_T$ , is on the order of

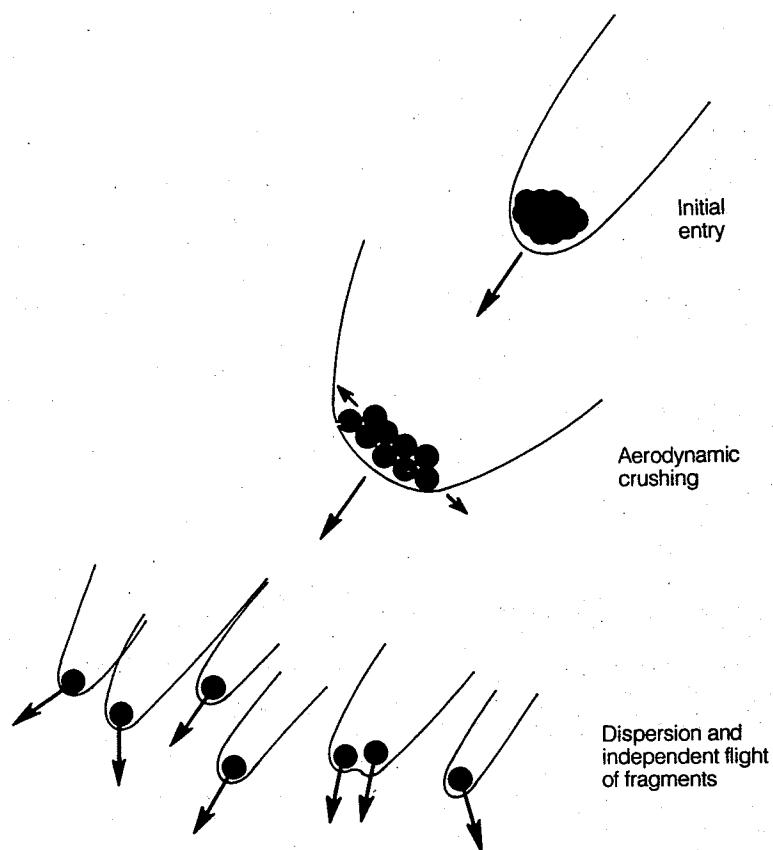
$$v_T \approx (\rho_g / \rho_p)^{1/2} v \quad (11.2.10)$$

The cross-range separation of the fragments,  $\Delta Y$ , is  $v_T t_f$ , where  $t_f$  is the flight time between breakup and impact. If breakup occurs at altitude  $Z_b$ , then the separation is

$$\Delta Y \approx (\rho_g / \rho_p)^{1/2} Z_b / \sin \theta \quad (11.2.11)$$

The cross-range separation in this equation reaches a maximum for breakup at two scale heights,  $Z_b = 2H$ , or about 15 km on earth and 30 km on Venus. For entry angles  $\theta$  near 15 degrees, the maximum cross-range dispersion on earth is predicted to be about 1 km, close to values observed in crater fields. On Venus the maximum dispersion is about 15 km, whereas on Mars, in the unlikely case that breakup occurs at two scale heights, the maximum dispersion is only 20 m.

It is noteworthy that the maximum cross-range dispersion in Equation 11.2.11 is independent both of meteoroid velocity and size. The size of the scatter ellipse on the ground is thus independent of meteoroid size. If the meteoroid is small, craters are scattered sparsely within the ellipse. If the meteoroid is larger, the resulting craters are



**Fig. 11.3** Schematic illustration of the atmospheric entry and breakup of a large meteoroid. Depending upon the meteoroid's size it may strike the ground during any of the stages shown here. If it strikes as a crushed but closely grouped collection of fragments, it produces a single crater that may, however, have an anomalous morphology. If it strikes as a collection of separate fragments, a crater field is the result. Small meteoroids burn up entirely before they strike the ground.

larger and occupy a larger fraction of the scatter ellipse's area. It is clear that when the meteoroid reaches some sufficiently large size the fragments will produce a crater that fills the entire scatter ellipse and only one crater will be observed, even though it may have been excavated by the nearly simultaneous impact of many separate fragments. On the earth this critical size corresponds to meteoroids with masses from about  $10^5$  to  $10^{10}$  kg, depending upon angle of entry. On Venus this range is  $10^8$  to  $10^{15}$  kg. Figure 11.4 illustrates the fate of iron meteoroids entering the atmospheres of earth and Venus as a function of size and entry velocity. Dashed contours indicate the diameter of the largest crater produced in the field. The meteoroid that produced the Henbury crater field was evidently not large enough to fill the scatter ellipse (Fig. 11.2), but some of the larger craters do overlap, illustrating the beginning of the process.

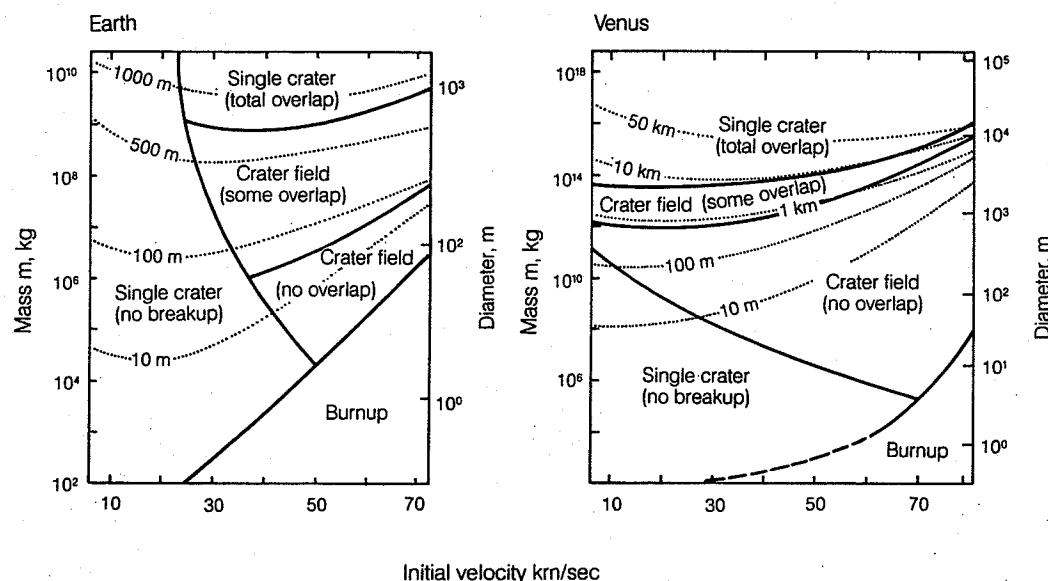
#### 11.2.4 Effect of breakup on crater morphology

It should not be surprising that a crater produced by a cluster of fragments from a crushed meteo-

roid differs morphologically and structurally from a crater produced by a solid meteoroid. Figure 11.5 illustrates two laboratory-scale craters produced by clustered impactors (basically small shotgun blasts). Although both possess the general crater morphology with raised rims and exterior ejecta blankets, they are anomalously shallow compared with craters produced by solid impactors and have nearly flat interior floors. Somewhat tighter clusters are observed to produce craters with central mounds. Small terrestrial craters such as Flynn Creek (3.5 km in diameter) and Decaturville (ca. 3 km in diameter) exhibit a similar morphology with broad, shallow excavation cavities surrounding a central region of chaotically broken rocks. This central region may be the scar of the largest fragment that reached the surface, whereas the shallow cavity is excavated by smaller fragments falling around it.

As meteoroid size increases, the importance of breakup and dispersion of the fragments progressively decreases. For example, the crater produced by a fragmented 10-km diameter meteoroid striking the earth would not be significantly different from that produced by a solid object if

## IMPACT CRATERING



**Fig. 11.4** The fate of large iron meteoroids entering the atmospheres of earth and Venus as a function of mass (or diameter) and velocity. The entry angle is assumed to be 15 degrees and the meteoroids have a relatively large crushing strength of 500 MPa. The areas outlined by solid lines denote the final outcome of the entry, and the dotted contours indicate the approximate diameter of the largest final crater (if there is more than one). *Modified from Passey (1980).*

the fragments spread 1 km laterally before they struck the ground, because the breakup results in only a 10 percent change in shape. A convenient parameter for gauging the importance of shape distortion is the aspect ratio,  $\alpha$ , which is defined as the ratio between the largest and smallest diameters of the meteoroid, whose shape is approximated as a triaxial ellipsoid. The theory of breakup (Melosh, 1981) shows that the aspect ratio of a collection of fragments produced from an initially spherical meteoroid is, upon striking the surface,

$$\alpha = \{1 + [L_c/(L \sin \theta)]^2\}^{1/3} \quad (11.2.12)$$

where the critical diameter for breakup is

$$L_c = 2(\rho_0/\rho_p)^{1/2} H \quad (11.2.13)$$

This equation is valid if breakup occurs at an altitude greater than about  $3H$ . Otherwise, the aspect ratio is smaller (it is only 20 percent of that computed from Equation 11.2.12 if breakup occurs at one scale height  $H$ ). Values of the critical diameter for Venus and earth are given in Table 11.2.

A meteoroid of the critical diameter will, upon striking vertically, have spread out to attain an aspect ratio  $\alpha = 8$ . That is, it will have spread laterally until its width is eight times its thickness. The meteoroid's shape is that of a fat pancake as it impacts the surface. Such a collection of fragments probably does not produce the same kind of crater that a solid impactor would. In fact, such a large aspect ratio may signal the incipient separation of the fragments into independently trav-

TABLE 11.2 Critical diameter for aerodynamic dispersion of meteoroids\*

Planet	Scale Height (km)	Atmospheric Density, $\rho_0$ ( $kg/m^3$ )	Critical Diameter $L_c$ (m)		
			Ice	Stony	Iron
Venus	14.9	70.1	7700	4400	2700
Earth	8.0	1.30	580	330	200
Mars†	10.6	0.0104	66	38	24
Titan	19.4	5.4	2900	1700	1000

\*Computed from Equation 11.2.13.

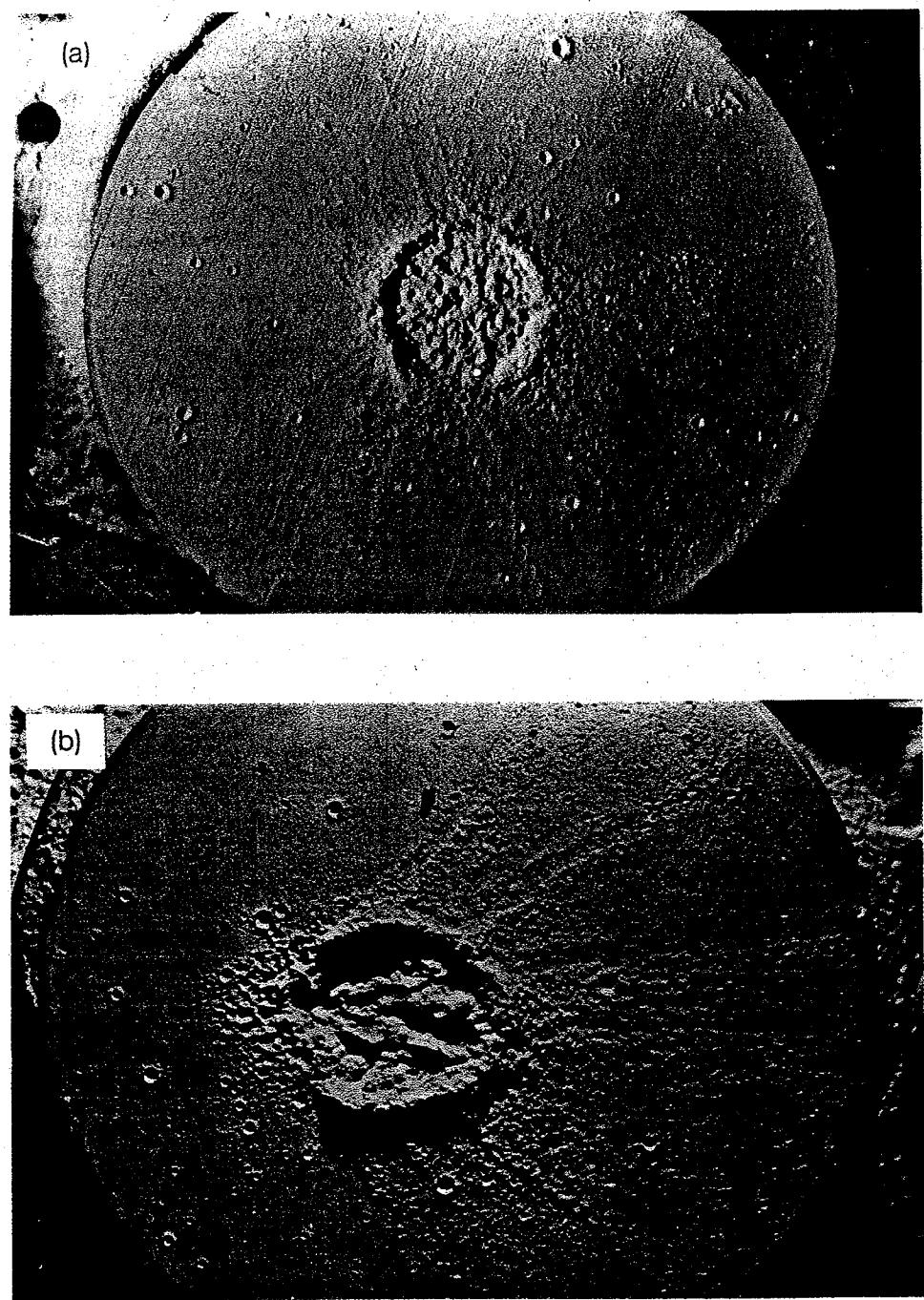
†These diameters assume that breakup occurs above three scale heights, which is highly unlikely on Mars.

eling bodies. If we adopt the criterion that the aspect ratio must be 2 or less for the final crater to approximate that produced by a solid projectile, then the projectile diameter must be roughly twice as large as the critical diameter  $L_c$  of Equation 11.2.13. Thus, a stony terrestrial impactor must be larger than about 700 m in diameter (and at near-vertical incidence) before the crater it produces can be considered comparable to craters produced on airless bodies. A projectile of this diameter produces a crater about 10 km in diameter. On Venus the atmosphere is even more important: the impactor must be about 9 km in diameter, corresponding to a crater nearly 80-km wide.

In summary, the presence of a dense atmosphere profoundly modifies the process of impact cratering. Small meteoroids may either be entirely prevented from striking the surface or strike at greatly reduced velocity. Even much larger me-

Fi  
or  
su  
(1

te  
m  
su  
th  
sp  
ai



**Fig. 11.5** Craters produced by vertical (a) and 45-degree oblique (b) impacts of clustered impactors on compacted pumice targets. Note the flat-floored interior morphology and prominent radial and subradial spokes in the ejecta. The craters are each about 7 cm in diameter. *From Schultz and Gault (1985), photo courtesy P. H. Schultz.*

eteoroids, however, are crushed during their atmospheric passage and, when they do impact the surface at high velocity, produce a crater form that has no analog on airless bodies. Only the very largest meteoroids can brush aside the atmosphere and produce craters that are rigorously comparable to those observed on airless planets and satellites.

### 11.3 VAPOR AND EJECTA INTERACTION WITH AN ATMOSPHERE

A new set of processes comes into play after a meteoroid succeeds in penetrating a planet's atmosphere and striking its surface. As the crater grows, the first material to leave the impact site is generally a plume of vaporized meteoroid and tar-

get material (Sec. 5.3), unless, of course, the impact velocity is too low for vaporization to be important (that is, less than about 15 km/second; see Secs. 5.2.3 and 7.10). If a vapor plume does not form, the expansion of the ejecta curtain is retarded by atmospheric resistance and the extent of the final ejecta blanket may be restricted. This topic was discussed in Section 6.3.5 in the context of Martian ejecta deposits and is illustrated in Figure 6.7. I will not repeat the discussion here, but merely refer the reader back to it. It is important to note, however, that in addition to modifying the ejecta blanket, the atmosphere winnows out the finer fraction of debris from the ejecta. This fine material may remain suspended for long periods of time (months to almost a year in earth's atmosphere). Such suspended dust may have important climatic consequences, which are discussed in the next chapter.

When a substantial vapor plume forms in a high velocity impact, it is the fastest and hottest material to leave the impact site, generally outrunning even the fastest solid ejecta. The ambient atmosphere is swept aside by this cloud of hot vapor, which has expanded from an initial pressure of many GPa. A shock wave forms where they meet and subsequently propagates away from the impact site as a strong airblast.

In relatively small impacts the vapor plume expands until its pressure equals that of the surrounding atmosphere. Approximating the expansion as adiabatic, the standard thermodynamic relation  $PV^\gamma = \text{const.}$  relates the gas pressure to the volume of the expanding cloud, where  $P$  is gas pressure,  $V$  is volume, and  $\gamma$  is the ratio of specific heats of the gas. The initial pressure  $P_i$  and volume  $V_i$  of the gas may either be obtained through simple approximations, discussed in Sections 4.5.1 and 5.2, or it may need to be computed numerically. If these two quantities are known, the radius of the equilibrated vapor cloud (or fireball, in military parlance)  $R_f$  is

$$R_f = [(3V_i/2\pi)(P_i/P_0)^{1/\gamma}]^{1/3} \quad (11.3.1)$$

where  $P_0$  is the pressure of the ambient atmosphere. Alternatively, if the energy  $E_a$  deposited in the atmosphere by the vapor plume is known,

$$R_f = 0.009E_a^{1/3} \quad (11.3.2)$$

Although the pressure in the expanded vapor cloud equals that of the surrounding atmosphere, the vapor is generally much hotter and so less dense than the ambient gas. The vapor cloud thus rises like a giant hot air balloon, driven upward by buoyancy forces (Fig. 11.6a). The center of the cloud rises faster than its periphery, causing the cloud to turn itself "inside out" and develop into a rising torus of hot gas. The timescale for torus formation is typical of gravitational stability processes, equal to  $(R_f/g)^{1/2}$ . The hot, rotating torus expands, cools, and mixes with the ambient air as

it rises. It eventually stabilizes at some level in the stratosphere and finally dissipates. The site of a small impact is thus the source of a classical mushroom cloud whose rise may also entrain dust and debris and carry them to high altitude.

The atmospheric phenomena associated with a very large impact, however, are quite different. If the computed radius of the equilibrated gas cloud  $R_f$  is larger than about one scale height  $H$ , then equilibrium is never reached. Instead, the expanding gases "blow out" of the top of the atmosphere into the relative vacuum above and continue expanding almost as if they had formed in a true vacuum. A large vapor plume expands over the site of the impact, and vapor (which may contain entrained ejecta) moves into the upper atmosphere at high velocity, restrained only by gravity. This process is illustrated schematically in Figure 11.6b. It has been proposed that this type of expansion is responsible for the ejection of tektites, glassy blobs of solidified impact melt, above the earth's atmosphere. Many tektite fields are strewn over distances of thousands of kilometers, distances that are hard to reconcile with anything but a long ballistic flight above most of the earth's atmosphere. Microtektites (glassy spherules  $\lesssim 1$  mm in diameter) form widely distributed layers in deep sea sediments, again posing a problem for ordinary atmospheric dispersal mechanisms.

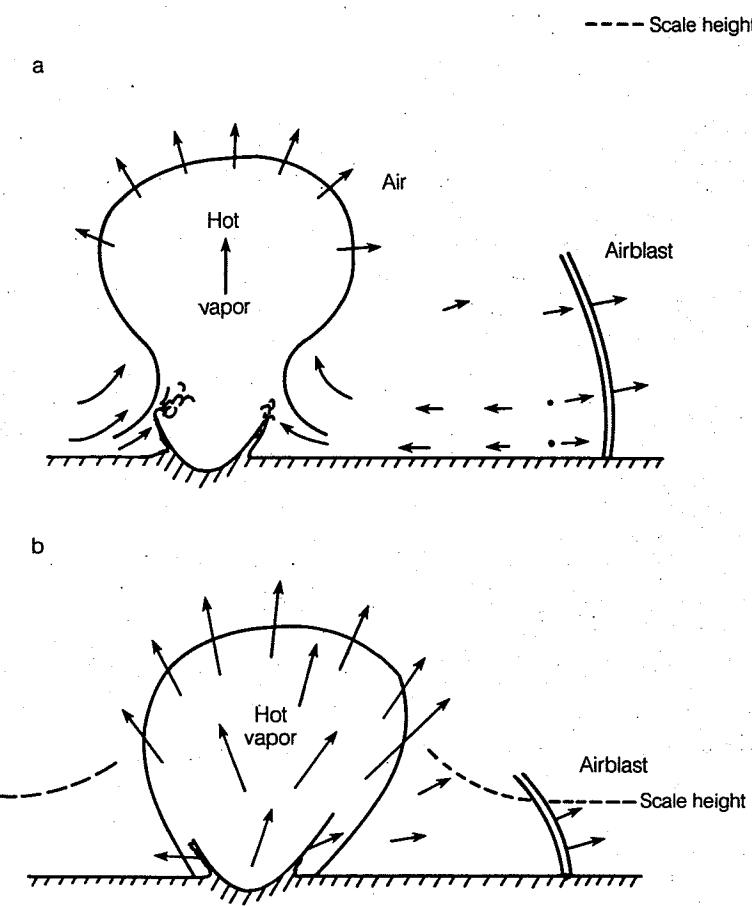
The conditions that determine the transition between fireball expansion and atmospheric "blowout" are not well known, especially for impacts. No nuclear detonation or known volcanic eruption has ever formed the "blowout" type of plume (which is fortunate from the point of view of dispersal of radioactive debris). The largest such detonation reported is a 1961 Soviet test in which about 60 MT ( $2.5 \times 10^{17}$  J) were deposited in the atmosphere. Numerical simulations suggest that a minimum of 150 MT ( $6.3 \times 10^{17}$  J) must be deposited in the atmosphere before "blowout" occurs. This corresponds to a 100-m-diameter stony meteoroid traveling at 30 km/second, which would create a 3-km diameter crater, assuming that 100 percent of the meteoroid's energy is deposited in the vapor plume. Since 100 percent conversion is unlikely, 3 km may be taken as the minimum size for a crater whose vapor plume could penetrate the atmosphere. It is noteworthy that the 22-km diameter Ries Crater in Germany is probably the source of the Moldavite tektites, which have been found 400 km from the parent crater, indicating that the Ries impact more than exceeded the transition energy. The 10.5-km diameter Bosumtwi Crater in Ghana is probably the source of the Ivory Coast tektites, so that the vapor plume of even a 10-km diameter crater may be sufficient to blow the atmosphere aside.

Our present understanding of vapor plume interactions with an atmosphere is not profound.

Fig.  
that  
The  
larg  
pan  
imp

Ho  
mo  
mo  
ma  
sm  
wel  
larg  
con  
am  
Cre  
pac  
pha  
her

FU  
Ge  
tha  
pe



**Fig. 11.6** Vapor plume interaction with the ambient atmosphere. In (a) the impact is small enough that a hot, buoyant bubble of gas forms after the vapor plume has equilibrated at ambient pressure. The gas bubble subsequently rises into a classic mushroom cloud. The vapor plume produced by a larger impact (b) would, if equilibrated, be larger than the atmosphere's scale height, so instead it expands at high velocity into the upper atmosphere, throwing gas and debris great distances from the impact site. Both large and small impacts generate strong airblasts in the lower atmosphere.

However, recent interest in impact-induced atmospheric loss (Sec. 12.9) has turned numerical modelers' attention to the problem and progress may be forthcoming. Although the interaction of small meteors with the atmosphere is an old and well-studied topic, work on the interactions of large meteoroids has just begun, much of it in the context of the effects of the putative 10-km-diameter asteroid or comet whose impact ended the Cretaceous era (Sec. 12.10). Future work on impact cratering will thus probably place more emphasis on atmospheric interactions than heretofore.

#### FURTHER READING

##### *General References*

There are no comprehensive treatments of the interactions between meteoroids and atmospheres that provide an adequate exploration of what happens after the meteoroid strikes the ground. The

reader interested in these phenomena is advised to seek what enlightenment may be had from the nuclear explosion literature. The most comprehensive and readily available treatment of this kind is

S. Glasstone (ed.), *The Effects of Nuclear Weapons*, U.S. Government Printing Office, Washington, D.C., 730 pp. (1962).

If you can find it, the first (1950) edition of this handbook offers a much better explanation of the physics of a nuclear explosion than the later editions, although it is less complete.

##### *Special Topics*

The flight of meteors through the earth's atmosphere has been the subject of many books and papers. A comprehensive treatment that is not too obsolete is

D. W. R. McKinley, *Meteor Science and Engineering*, McGraw-Hill, New York, 309 pp. (1961).

An excellent and more recent treatment of the ablation and breakup of single meteoroids is

B. Baldwin and Y. Sheaffer, Ablation and entry of large meteoroids during atmospheric entry, *J. Geophys. Res.* 76: 4653-4668 (1971).

A fine recent discussion of the Tunguska event and its atmospheric effects can be found in

R. P. Turco, O. B. Toon, C. Park, R. C. Whitten, J. B. Pollack, and P. Noerdlinger, An analysis of the physical, chemical, optical and historical impacts of the 1908 Tunguska meteor fall, *Icarus* 50: 1-52 (1982).

Atmospheric breakup and crater field formation is treated by

Q. R. Passey and H. J. Melosh, Effects of atmospheric breakup on crater field formation, *Icarus* 42: 211-233 (1980).

The special crater forms produced by clustered impactors were studied experimentally by

P. H. Schultz and D. E. Gault, Clustered impacts: Experiments and implications, *J. Geophys. Res.* 90: 3701-3732 (1985).

The crushing and distortion of large projectiles is the subject of

H. J. Melosh, Atmospheric breakup of terrestrial impactors, in P. H. Schultz, R. B. Merrill (eds.), *Multiring Basins, Proc. Lunar Planet. Sci.* 12A, pp. 29-35 (1981).

A good discussion of the interaction between the atmosphere and an expanding vapor plume from an impact or explosion, with some background on nuclear and volcanic fireballs is in

E. M. Jones and J. W. Kodis, Atmospheric effects of large body impacts: The first few minutes, *Geol. Soc. Amer. Special Paper* 190: 175-186 (1982).

The G.S.A. Special Paper 190 referred to above is a unique collection of papers that cover many other topics relating to both the physical and biological effects of a large impact on the earth. In particular, the possible consequences of a large impact in the ocean are well discussed by

D. E. Gault and C. P. Sonett, Laboratory simulation of pelagic asteroid impact: Atmospheric injection, benthic topography, and the surface wave radiation field, *Geol. Soc. Amer. Special Paper* 190: 69-92 (1982).

## XII. IMPACT CRATERING AND PLANETARY EVOLUTION

### 12.1 INTRODUCTION

Over the last few decades, it has become apparent that impact cratering has played a major role in the formation and subsequent history of the terrestrial planets and satellites. It has long been recognized that impacts were important during the growth of the planets from planetesimals, but only recently have the effects of major impacts on the earth's climate and biological evolution been considered. Still more recent work suggests that the moon was formed by a truly great impact in which a Mars-size protoplanet struck the protoearth, ejecting vapor and debris that eventually condensed into the moon.

Aside from their scientific interest, impact craters have also attained modest economic importance. The Sudbury structure in Ontario, Canada, which contains one of the world's largest deposits of nickel and other valuable metals, is now recognized to be a tectonically distorted impact crater that was initially about 140 km in diameter. Oil has been produced from buried craters in sedimentary basins. Two examples are the 10-km-diameter Red Wing Creek Crater in the Williston Basin and the 3.2-km diameter Newporte structure in North Dakota, which formed in basement rocks. In both these craters, rocks shattered by the impact produce a buried reservoir with high porosity and permeability, conditions ideal for the accumulation and exploitation of hydrocarbons. The Manicouagan Crater in eastern Quebec has been put to still another economic use. A 60-km-diameter annular depression surrounding the crater's center was scoured out by the Laurentide ice sheet and is currently employed as a water reservoir.

In this chapter I will briefly explore the major ways in which impact cratering has affected the evolution of the planets and the history of the earth. Many of these topics have large literatures of their own, to which references are given at the end of the chapter.

### 12.2 ACCRETION OF THE PLANETS

Modern theories suggest that the sun and the planets formed together about  $4.6 \times 10^9$  years ago from an interstellar cloud that had become unstable against gravitational collapse. The details of this collapse are beyond the scope of this book; however, the result was a protosun surrounded by

a much lower mass nebula composed of gas and condensed dust. Gas-dynamic interactions concentrated the dust into a thin disk in the sun's equatorial plane. This disk itself became gravitationally unstable and broke up into clumps that eventually aggregated together to form planetesimals with diameters ranging from 1 to about 10 km. At some stage during this early accretion, most of the gas was lost from the terrestrial planet zone. From that time forward, impact dominated the interactions of the planetesimals and mediated their growth into the planets we recognize today.

Impact velocities during the early stages of planetesimal accretion must have been so low that the typical outcome of an impact between two planetesimals was a larger planetesimal rather than many small fragments. Although the conditions under which accretion rather than disruption occurs are not well known, much work in the past has focused on a dimensionless ratio dubbed the "Safronov number"  $\Theta$  by G. W. Wetherill (1980)

$$\Theta = v_{\text{esc}}^2 / 2v^2 \quad (12.2.1)$$

where  $v_{\text{esc}}$  is the minimum collision velocity of the two objects with masses  $m_1$ ,  $m_2$ , and radii  $R_1$  and  $R_2$ , respectively:

$$v_{\text{esc}} = [2G(m_1 + m_2)/(R_1 + R_2)]^{1/2} \quad (12.2.2)$$

where  $G$  is Newton's gravitational constant, and  $v$  is the relative velocity prior to a close encounter. This minimum collision velocity is essentially the escape velocity of the bigger planetesimal when one of the impacting bodies is much larger than the other. The actual impact velocity  $v_i$  is

$$v_i = (v^2 + v_{\text{esc}}^2)^{1/2} \quad (12.2.3)$$

Accretion is expected to occur in collisions for which the Safronov number is greater than approximately 1, whereas both planetesimals are disrupted in collisions when  $\Theta \lesssim 0.2$ . This criterion ignores the possibility that some form of material strength may hold the planetesimals together, the effects of rotation or oblique impact or the nature of the planetesimals' surfaces, whether rough, powdery, sticky, etc. Little is yet known about the collisions of either solid objects or loose aggregates at velocities sufficiently low to simulate the early phases of accretion, although efforts are underway to study these processes experimentally.

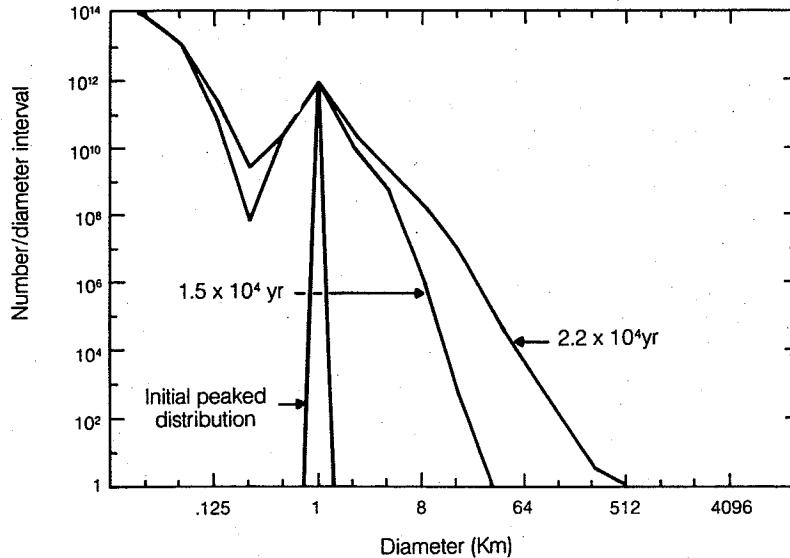
The original planetesimal population evolves as a result of collisions, with smaller bodies accumulating into ever-larger aggregations while some small bodies break up into still smaller fragments (see Fig. 12.1 for the outcome of one set of computations). The details of this growth, however, are highly controversial. In general, it seems that the average encounter velocity  $v$  rises as the mean size increases so that  $v$  is always comparable to the escape velocity  $v_{esc}$  of the largest object. This rise in encounter velocity is due to gravitational perturbations during near misses of the smaller bodies with the larger ones. As the mean encounter velocity rises, the collisions of small bodies with one another become destructive so that only the largest bodies continue to grow. At present it is not clear whether this process produces a small number of large planets directly or whether it results in a large number of subplanetary objects that have isolated themselves by sweeping up all the available material in their neighborhood during "runaway growth." In the latter case, other disturbances (Jovian perturbations?) are needed to cause orbits to intersect and permit further growth.

Recent work (Wetherill, 1985) suggests that the final stages of accumulation may be more violent than had previously been thought. At one extreme one can imagine the terrestrial planets growing by direct accumulation of small planetesimal-sized objects. In this case the final planets should have small eccentricities, inclinations, and obliquities and very long rotational periods (hundreds to thousands of hours). At the other extreme, Wetherill's (1985) computations show that a few thou-

sand subplanetary objects (masses  $10^{22}$  to  $10^{24}$  kg) evolve into a few (3 to 6) planets by a long sequence of major collisions over the course of  $10^7$  to  $10^8$  years. During this growth most bodies are struck several times by others approximately half their size. Large amounts of impact melting, devolatilization, and other poorly studied phenomena may occur during these violent events. Such collisions are close to disruptive and produce a final set of planets with significant eccentricities and inclinations and somewhat chaotic obliquities with relatively short rotational periods (tens of hours). Needless to say, this final state resembles our solar system more closely than the outcome of direct accretion of planetesimals. The study of impacts between planetary objects of comparable size is in its infancy but promises to yield many insights into the history of the solar system.

### 12.3 THE THERMAL STATE OF GROWING PLANETS

Geologic study has shown that all of the terrestrial planets have been at least partially molten and have differentiated at some time in their histories. The eucrite meteorites show that even asteroid-size bodies were capable of producing basaltic melts during the early eras of solar system history. The source of heat for these events is elusive. Short-lived radioactive elements, such as  $^{26}\text{Al}$ , electromagnetic induction heating, impact heating, and the energy released by core formation in the larger planets have all been proposed as important contributors to the thermal budget of the



**Fig. 12.1** A numerical computation by Greenberg et al. (1978) demonstrated how an initial distribution of  $10^{12}$  1-km diameter planetesimals evolves into a population that contains objects as large as 500 km in diameter in addition to a large number of smaller fragments produced by destructive encounters. The timescale for this growth is only about  $10^4$  years.

early planets and asteroids. In the following section I will consider the major effects that impacts have on the thermal state of a growing planet.

The amount of heat deposited in a planet growing by impact accumulation of surrounding debris is a function of the size of the bodies striking the planet's surface. The rate at which heat is deposited is a fraction of the kinetic energy of the impacting material. This in turn is contributed partly by the nonzero encounter velocity  $v$  and partly by the gravitational energy gained as the material falls to the planet's surface. The total energy  $E_a$ , per unit mass, of accreted material striking the surface of a planet of mass  $M$  and radius  $R_p$  is

$$E_a = v^2/2 + GM/R_p \quad (12.3.1)$$

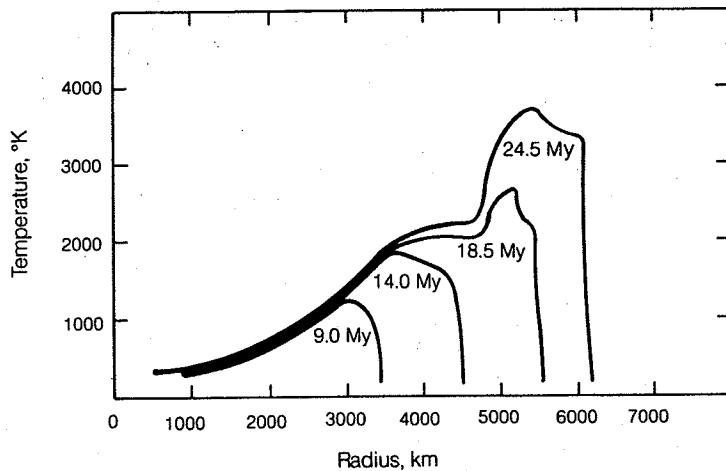
Of this energy a total of  $kE_a$  is available to heat the planet, where  $k$  is an efficiency factor that is close to 1 in most cases (some impact energy may be lost to the planet immediately after impact if significant amounts of ejecta escape the planet, thus reducing  $k$ ).

If most of the infalling material is in the form of relatively small particles, the accretion energy is deposited in a thin surface layer and most of this heat is radiated away by thermal radiation at a rate given by  $\epsilon\sigma(T_s^4 - T_0^4)$ , where  $\epsilon$  is the emissivity of the surface,  $\sigma$  is the Stefan-Boltzmann constant,  $5.669 \times 10^{-8} \text{ J/m}^2 \text{ K}^4 \text{ second}$ ,  $T_s$  is the surface temperature, and  $T_0$  is the radiation equilibrium temperature in the absence of impact heating. The remainder of the heat not radiated away is either conducted into the planet's interior or consumed in raising the temperature of the surface layer. Studies using thermal balance models of this kind require extremely fast accretion times for temperatures to exceed the melting point in any of the terrestrial planets. Thus, the earth would have to have grown to most of its final di-

ameter in  $10^3$  to  $10^4$  years for melts to have been generated by impact heating. Such rapid accretion times are inconsistent with accretion models and seem at first to suggest that impact heating is unimportant in growing planets.

However, if the infalling material is in the form of larger projectiles, objects 10 km or more in diameter, a significant fraction of the total energy is buried deeply enough below the surface that thermal radiation loss is not in equilibrium with the energy deposition rate. In this case much of the energy deposited goes into raising the temperature of the planet's interior. Although the exact fraction is uncertain, estimates suggest that it is as large as 50 percent of the accretional energy. This fraction is conventionally denoted  $h$ , so that the available thermal energy is a  $h k E_a$ , where  $h \approx 0.5$  in this projectile size range. Since  $E_a$  is roughly proportional to  $GM/R$  (Eq. 12.3.1; note that in accretional models  $v^2$  is comparable to  $GM/R$ ), and  $M$  is roughly proportional to  $R^3$ , the original temperature in the planet due to accretional heating is proportional to  $R^2$ . Thus, the outside of a growing planet is hotter than its interior, and heat is conducted downward. Convection, of course, is prohibited in such an inverted temperature gradient. Because the accretional energy is buried deeply enough to prohibit thermal radiation loss, and conduction is a slow process, melting and even vaporization due to accretional heating may occur in planets that grow over periods of  $10^6$  to  $10^7$  years (see Fig. 12.2).

Although this has not yet received much study, it appears that giant impacts of the type suggested by Wetherill's (1985) models produce a still different temperature distribution. During such a collision much of the impact energy is carried in a strong shock wave that propagates through the target planet. As the shock wave decays, this energy is deposited more deeply and uniformly than



**Fig. 12.2** Thermal evolution of a hypothetical earth growing by moderate-sized impacts. The temperature profile is inverted, with hot exterior and cool interior, thus inhibiting convection. The melting point of the earth's outer shells is reached after about 10 Myr of growth. After Kaula (1979).

predicted for 10-km diameter impactors and may thus lead to a hotter initial core. The implications of this model for such processes as core formation and crustal differentiation are presently unknown.

#### 12.4 PLANETARY EROSION BY IMPACTS

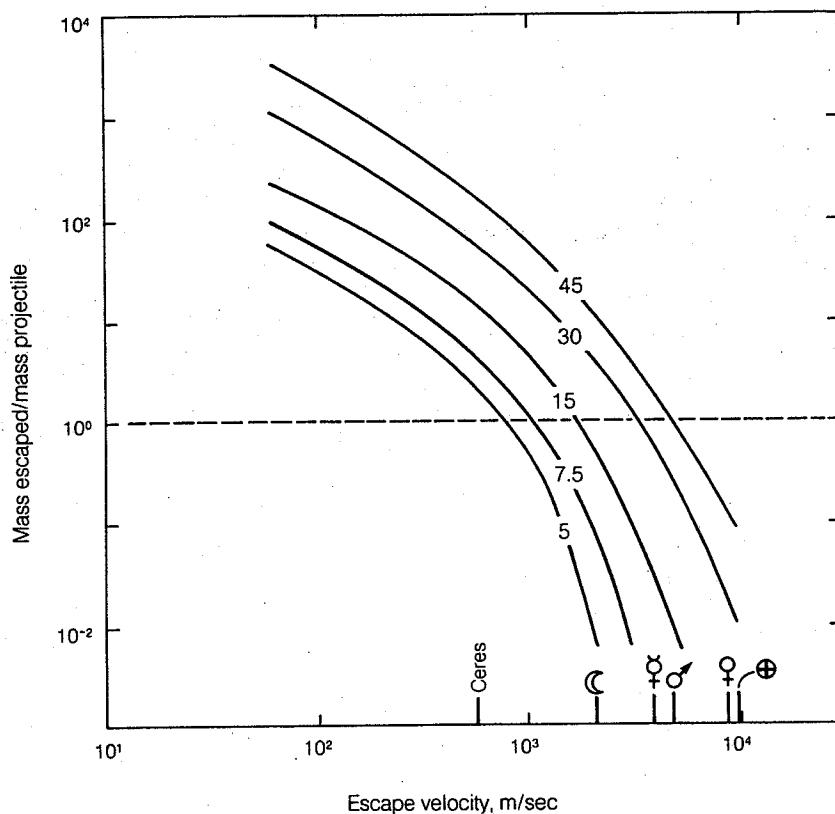
Planetary erosion is the inverse of accretion. Whereas low-velocity impactors striking a planet's surface may add to its total mass, high-velocity impacts may produce a net mass loss. The crossover velocity between these two regimes is a function of the planet's escape velocity and, to a lesser extent, of the impactor's composition. Figure 12.3 illustrates the ratio between the mass ejected from a planet to the impactors mass as a function of escape velocity and the impact velocity. The curves in this figure were computed from a series of numerical studies of the vertical impact of an anorthosite projectile on a planar anorthosite target (O'Keefe and Ahrens, 1977). The crossover between accretion and erosion occurs when the ejected mass equals the impactor's mass. This crossover occurs at velocities of 20, 35, and 45 km/second for the moon, Mercury, and Mars, respectively. From the current mean encounter ve-

locities, it would appear that the moon is being eroded and Mars is still accreting at the present time.

Encounter velocities in the asteroid belt are currently in the vicinity of 5 km/second. The largest asteroid, Ceres, has an escape velocity of about 600 m/second, so, from Figure 12.3, it appears that Ceres and all of the other asteroids are currently being eroded by impacts. This conclusion is strengthened by the observation of numerous families of asteroids that are evidently fragments of a once-larger object. Most meteorites falling to the earth today have been exposed to cosmic rays in space for only a few tens of millions of years (stony meteorites) to a few hundreds of millions of years (iron meteorites). These meteorites are evidently fragments of larger bodies that have been ejected in geologically recent time. The most plausible source for most meteorites is thus impact ejection from asteroids.

#### 12.5 CATASTROPHIC DISRUPTION OF ASTEROIDS AND PLANETESIMALS

Impact erosion occurs as many small impacts wear away the surface of the bombarded planet or



**Fig. 12.3** Mass ejected from a planet by a meteoroid impact as a function of the planet's escape velocity and the impact velocity. These curves are based on a computation that assumes a gabbroic anorthosite projectile striking a target of similar composition. The target planet's radius of curvature is not taken into account in these computations. *After a figure in O'Keefe and Ahrens (1977). Copyright 1977 by the AAAS.*

asteroid. Sufficiently large impacts may, however, disrupt the target and break it into fragments much smaller than the original body in a single catastrophic event. This process has received extensive laboratory study at impact velocities of a few km/second and for targets in the centimeter to meter size range. The outcomes of these impacts can be summarized by a dimensionless number  $f_i$ , the ratio between the mass of the largest remaining fragment to the mass of the target before impact. Thus, an impact with  $f_i \approx 1$  is erosive rather than disruptive because the target remains largely intact. If  $f_i \approx \frac{1}{2}$  the target has suffered a major breakup, whereas if  $f_i \ll 1$  the target is dispersed into small pieces.

In work to date,  $f_i$  is assumed to be a function solely of some measure of the target's strength and the kinetic energy imparted by the projectile,  $E_p = \frac{1}{2}m_p v_i^2$ . Fits to the data suggest relations of the form

$$f_i = A E_p^{-d} \quad (12.5.1)$$

Where  $A$  is a dimensional constant for a given target material and the slope  $d$  ranges between 0.8 and 1.25 for targets of ice, basalt, and granite. Another way of writing this expression is to note that  $A$  has dimensions (energy) $^d$  so that

$$f_i = \alpha (E_{\text{disrup}}/E_p)^d \quad (12.5.2)$$

where  $\alpha$  is a dimensionless constant of order one and  $E_{\text{disrup}}$  is the "disruption energy" of the target.  $E_{\text{disrup}}$  is of order  $Y_t V_t$ , where  $Y_t$  is the target's tensile strength (since most fractures produced by impacts are tensile spalls) and  $V_t$  is the volume of the target. As expected intuitively, for fixed strength larger targets require more energy to disrupt than small ones and for fixed target volume weak targets are more easily disrupted than strong ones.

Recent work suggests that the functional form of Equation 12.5.1 is not adequate to describe the outcome of an impact. The data can be fit better if some dependence of  $f_i$  on strain rate is permitted. This is a natural consequence of the dependence of  $E_{\text{disrup}}$  on  $Y_t$ , since the dynamic tensile strength is known to depend upon a fractional power of the strain rate. For the purposes of impact experiments, the strain rate is roughly the inverse of the rise time of the stress pulse  $\tau_s$ , given by Equation 5.2.1 as  $L/2v_i$ , where  $L$  is the projectile diameter and  $v_i$  is the impact velocity.

The gravitational binding energy of the target is negligible in laboratory experiments. However, it may be significant for asteroids and is dominant for planetary-size bodies. Although there is little direct information on catastrophic disruption when both gravity and strength are important, the form of Equation 12.5.2 suggests that both may be roughly taken into account if the disruption energy is the sum of both the work required to frac-

ture the target,  $Y_t V_t$ , and the gravitational binding energy of the target,  $\frac{1}{2} G m_t^2/R_t$  for a uniform spherical target of mass  $m_t$  and radius  $R_t$ :

$$E_{\text{disrup}} \sim a Y_t V_t + b G m_t^2/R_t \quad (12.5.3)$$

where  $a$  and  $b$  are dimensionless constants of order one. If the strength  $Y_t$  is negligible, Equation 12.5.2 with the definition of Equation 12.5.3 essentially reduces to the fragmentation criterion for planetesimals discussed in Section 12.2. Strength and gravitational binding energy make approximately equal contributions to the disruption energy when

$$R_t \approx (Y_t/G\rho_t^2)^{1/2} \quad (12.5.4)$$

where  $\rho_t$  is the target's density. For a silicate asteroid with the relatively large strength of 0.1 GPa this crossover takes place at a radius of about 400 km, assuming  $a$  and  $b$  in Equation 12.5.3 are roughly equal. Other estimates place the transition radius at about 100 km. In either case, strength is likely to be a factor in the disruption of most asteroids but is unimportant for planetary-size objects.

## 12.6 IMPACT-INDUCED SEISMIC MODIFICATION OF TOPOGRAPHY

The impact of a large meteorite on a planet may modify adjacent features by a strong seismic shaking, even if the impact is not large enough to actually disrupt the planet. The seismic-energy equivalent of an impact was derived in Section 5.2.6, where it was shown that an impact large enough to create Meteor Crater, Arizona, radiates an amount of seismic energy comparable to a magnitude 5.6 earthquake. An impact of the size of the one that created the lunar crater Copernicus corresponds to a magnitude 10 earthquake.

In spite of these theoretical predictions, few surface features on the moon or other planets can be directly attributed to impact-induced seismic shaking. The problem is that, as in terrestrial earthquakes, most of the seismic energy is at such long wavelengths that adjacent slopes are accelerated up and down as a whole, so that only small differential accelerations of the bedrock and overlying rock debris develop to cause landslides.

The largest cratering events, however, such as Imbrium Basin on the moon or Caloris on Mercury, impart so much energy to their target planet that some sort of major modification might be expected. Shortly after the first encounter of Mariner 10 with Mercury in 1974 it was recognized that an area of very unusual topography (informally dubbed "weird terrain," but since officially known as "hilly and lineated terrain") is confined to the antipodal region of the Caloris impact structure (Fig. 12.4). The ruined crater rims and intersecting straight valleys trending in directions

of a  
small  
and met  
N  
imp  
tai  
may  
vert  
pla  
and exte  
anti  
sup  
may  
thei

12.  
One  
erai  
can  
in 1  
ide:  
fars  
ma  
nisi  
whi  
era  
rec  
Ins  
the  
ma  
the  
lun  
stre  
mu  
cer  
sou  
im  
ma  
an  
at 1  
ter  
abo  
7.1  
ca  
ing  
oc  
ma  
he  
be

no  
in

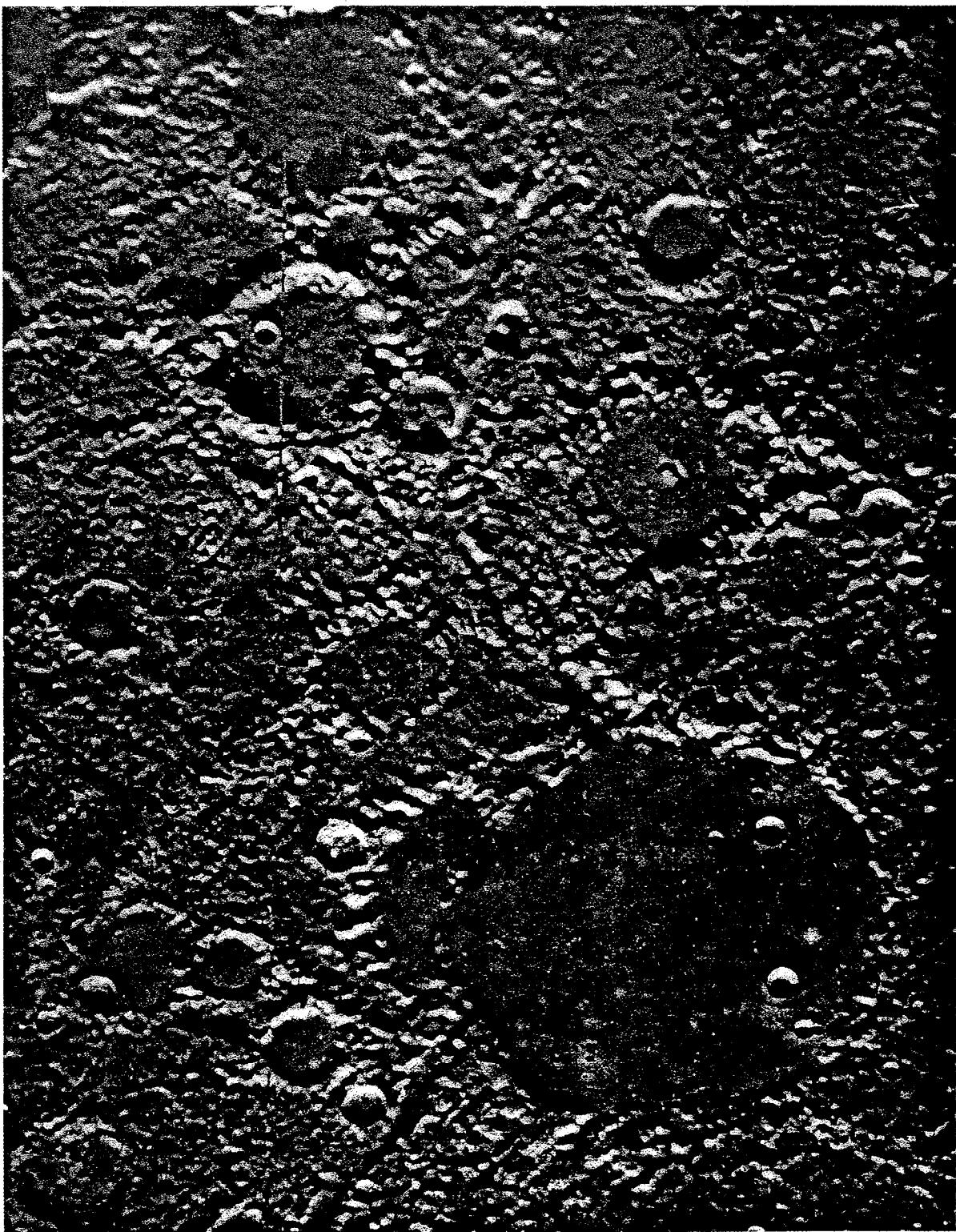


Fig. 12.4 A section of the "hilly and lined" terrain on Mercury antipodal to the 1300-km diameter Caloris Basin. The criss-crossing troughs and irregular crater rims are believed to be the result of seismic shaking produced by the Caloris impact. The picture is 543 km across. *Mariner 10 FDS 27370.*

of apparent structural weakness suggest that the terrain may have been modified by strong seismic shaking that jostled adjacent crustal blocks up and down relative distances measured in kilometers.

Numerical simulations of the effect of a large impact on a planet demonstrate that under certain conditions the seismic energy of the impact may be focused at its antipode and that general vertical oscillations of many kilometers may take place there. Subsequent to the discovery of hilly and lineated terrain on Mercury, similar (but less extensive) terrains have been found on the moon antipodal to the Imbrium and Orientale basins, supporting the idea that the very largest impacts may have important geomorphic effects even at their antipodes.

### 12.7 IMPACT-INDUCED VOLCANISM

One of the recurring themes of the older lunar literature was the idea that impacts may induce volcanic activity as a result of the heat they deposit in the target planet's crust. The validity of this idea was questioned when photos of the moon's farside revealed a number of large basins that lack mare basalt infillings. The impact-induced volcanism idea was dealt an even more severe blow when dating of samples returned from the moon showed that mare volcanism postdated basin excavation by as much as 1000 Myr. It is now generally accepted that mare volcanism had little directly to do with the basin-forming impacts. Instead, the basins provided low spots in which the mare basalts accumulated. Farside basins lack mare fill primarily because the crust is thicker on the farside and the basins lie higher above the lunar geoid than the nearside basins. At most, stresses or fractures associated with the lunar multiring basins have provided avenues for ascent of magma that developed from internal heat sources independent of the basin-forming impacts.

The geometry of the excavation flow (Sec. 5.5) makes it clear that most of the melt produced by an impact is within the excavation flow zone and at most coats the surface of the final collapsed crater. Only in the very largest impacts (greater than about 1800 km in diameter on the moon; see Sec. 7.10) does the melt volume exceed the volume excavated. Large impacts may even have a net cooling effect on the planet they strike: when collapse occurs the stratigraphic uplift beneath the crater may lift isotherms closer to the surface so that heat originally buried deep within the planet may be more readily conducted to the surface.

The pressure relief induced by this stratigraphic uplift is seldom sufficient to cause melting, especially on the moon, where the pressure gradient is not large anyway. Although pressure-relief melting may be important beneath the earth's mid-

ocean ridges, where convection currents lift hot mantle material on the order of 1000 km, even the largest impacts on the moon have lifted material less than 100 km, corresponding to a pressure relief of only 0.5 GPa. The pressure slope of silicate compositions likely to be of interest in the moon corresponds to only about 0.6 K/km, so a 100-km uplift would bring a mass of rock only 60K closer to the melting point (and this neglects the latent heat of melting, which, if included, would limit the amount of melt to a tiny fraction of the rock, even if it started out at the melting temperature before uplift). Pressure-relief melting that is due to impacts is thus unlikely to be of any importance on the moon, unless the target rock was already partially molten to begin with. Even on the earth, pressure-relief melting requires unusually large thermal gradients to begin with.

It is probably safe to say that, to date, there is no firm evidence that impacts can induce volcanic activity. Both the Sudbury and Vredefort structures on earth, however, are associated with massive igneous intrusions whose genesis has been supposed to be directly associated with the impact, although even in these cases it may be that the impact only provided an avenue for melts from the lower crust to pierce through the surface rocks. Further study should be able to resolve this issue.

### 12.8 ARCHEAN GEOLOGY OF THE EARTH

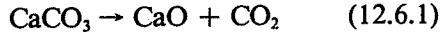
Evidence from the moon suggests that it underwent a period of especially heavy meteoritic bombardment before  $3.8 \times 10^9$  years ago. The oldest known rocks on earth are of comparable age, and many archean geologists have speculated that some of the unique features of rocks formed during that era may be related to large meteoritic impacts. Estimates based on the lunar crater record show that more than 100 impact structures with diameters greater than 1000 km should have formed on earth (Grieve, 1987). Although the subject of large archean impact structures has provided a fine topic for speculation, so far only one piece of direct evidence for such structures has been found.

D. R. Lowe and G. R. Byerly (1986) have recently reported the discovery of 30-cm- to 2-m-thick beds of sand-sized silicate spherules from 3.2 to 3.5 Gyr Archean greenstone belts in both South Africa and Western Australia. These spherule beds are best interpreted as melt ejecta from large impact events, and are the first direct evidence for an enhanced impact flux in the earth's early history. The beds extend over at least 100 km laterally, and are associated with elevated levels of iridium and shock features, making any other origin than impact seem highly unlikely. Hopefully, the wider awareness of the effects of

impact that has recently developed within the geologic community will lead to the discovery of more evidence for large Archean impacts in the future.

### 12.9 IMPACTS AND THE ORIGIN OF PLANETARY ATMOSPHERES

In the terminal phases of planetary accretion, infalling material strikes the surface of the planet at high enough speed to melt or vaporize most of the impactor. This accreted material may carry substantial inventories of volatiles, which are then dispersed over the growing planet by the vapor plume that subsequently expands out of the crater. Of particular interest is the decarbonation reaction



which begins at about 45 GPa and is substantially complete at about 70 GPa. This reaction may liberate carbon dioxide directly into the atmosphere of the target planet. That carbon dioxide, along with other released volatiles such as water and inert gases, may form an "instant atmosphere" far out of chemical equilibrium with the rest of the planet.

The study of such impact-generated atmospheres is just beginning, but preliminary results suggest that impact may be a major factor in the early atmospheres of the terrestrial planets. Future work will have to take account of the tendency of gases released by the impact to recombine with ejecta on the surface of the planet and the possible burial of this material by later ejecta blankets.

One interesting suggestion is that large impacts may be capable of ejecting a significant portion of the atmosphere, primarily by interaction between the ambient atmosphere and fast ejecta. This process, if it is important, may help to resolve some of the puzzling differences among the rare gas inventories of Venus, the earth, and Mars.

The earth's present atmospheric composition of 77 percent N<sub>2</sub> and 21 percent O<sub>2</sub>, which is primarily of biological origin, presents a unique situation for alteration by impact events. It is well known that lightning and nuclear explosions can induce a reaction between nitrogen and oxygen to produce NO:



Shock compression of the atmosphere by the penetration of a fast meteoroid can also induce this reaction. The 1908 Tunguska meteor, which flattened large areas of dense Siberian forest, also produced an estimated  $3 \times 10^7$  tonnes of NO (Turco et al., 1982). Subsequent to the impact the surrounding forest underwent a period of unprecedented growth, presumably due to the reaction

of NO with O<sub>2</sub> to form NO<sub>x</sub>, which fell to the ground and fertilized the region.

The amount of NO produced in the earth's atmosphere by meteors can be expressed in terms of the energy deposited. This may be nearly the entire kinetic energy of small meteoroids that are stopped by the atmosphere, or a fraction (Eq. 11.2.9) of the energy of larger meteoroids. Meteoroids are believed to be relatively efficient producers of NO, generating approximately one molecule of NO for each 40 eV deposited in the atmosphere, or about  $7 \times 10^{-9}$  kg of NO per Joule.

### 12.10 BIOLOGICAL EXTINCTIONS AT THE END OF THE CRETACEOUS

The cause of the widespread extinctions at the end of the Cretaceous Era, some 65 Myr ago, had long been a source of nearly unconstrained scientific and popular speculation until 1980 when a paper by L. W. Alvarez, his son Walter, F. Asaro, and H. V. Michel appeared in the journal *Science*. This paper reported the discovery of anomalously large abundances of the element iridium in a thin clay layer that marked the transition between the Cretaceous and Tertiary eras in a succession of marine sediments near Gubbio, Italy. Iridium is very scarce on the earth's surface (because it is siderophile, most of the earth's iridium inventory is probably in the core), but its abundance is much higher in undifferentiated meteorites. Alvarez et al. interpreted the results as indicating that a 10-km diameter asteroid had struck the earth, scattering the iridium over the globe. The physical and climatic effects of this great impact seemed to be more than adequate to account for the extinctions, which obliterated all land animals with a body mass greater than about 10 kg and nearly exterminated the marine calcareous nanoplankton and planktonic foraminifera. It is estimated that about half the genera extant perished in the event. On the other hand, land plants and many invertebrates were scarcely affected.

Although the reality of a large impact at the end of the Cretaceous continues to be argued, an impressive amount of evidence has been found to support the hypothesis. Iridium anomalies at the Cretaceous-Tertiary boundary have been found at 75 sites around the world, in both marine and terrestrial sediments. Other siderophile elements such as gold, osmium, and platinum are associated with the iridium in approximately chondritic ratios. In addition, sanidine spherules, which may be equivalent to microtektites, have been found in many locations. Perhaps the most telling evidence is the discovery of shocked quartz grains (up to the size of fine sand) in the boundary clay at widely separated localities. There is no natural process other than impact that is capable of pro-

ducing these shock features. The wide distribution of the shocked quartz argues for a large event. Recently, large concentrations of carbon soot have been found in boundary clays, suggesting that the impact may also have ignited global wildfires.

The only piece of evidence lacking to prove the case for a large impact is the crater itself. In spite of much effort neither the crater nor even any clues to its general location have yet been found. An impact of the required size should have produced a crater 100 km or more in diameter, and such a crater should not be difficult to find. One possibility is that the impact occurred in an ocean (which is, of course, more probable than an impact on land) in an area where the seafloor has been subducted since the Cretaceous. Although this solution is not very satisfying intellectually, it actually has a fairly high probability because roughly half the seafloor extant at the end of the Cretaceous has subsequently been subducted. In the meantime many geologists are diligently searching for some clue to the whereabouts of the crater. Another possibility, suggested by evidence that the Cretaceous-Tertiary extinctions may have been protracted over centuries, is that not one, but many, impacts occurred in rapid sequence, no single event being large enough to produce an unusually large crater.

The actual mechanism by which so many forms of life were extinguished is complicated and may never be fully known. The physical, climatological, and biological effects of an asteroidal impact are examined by many authors in a remarkable collection of papers published shortly after the Alvarez et al. discovery. This volume, Geological Society of America Special Paper 190, *Geological Implications of Impacts of Large Asteroids and Comets on the Earth*, contains detailed accounts of the mechanisms I will briefly summarize here.

One of the first ideas was that dust raised by the impact would be globally distributed. Dense dust in the atmosphere would screen out sunlight, bring photosynthesis to a halt, and disrupt both marine and terrestrial food chains. Most animals could not survive an extended period of darkness but plants, by virtue of their dormant seeds, could reestablish themselves when the dust cleared. Detailed studies of the behavior of a dense dust cover (Toon et al., 1982) showed that large quantities of dust could remain suspended only 3 to 6 months, by which time coagulation of the dust grains would have increased their sizes and allowed them to fall out of the atmosphere. These authors find that light levels on earth's surface would be too low for vision for 1 to 6 months after the impact and too low for photosynthesis for 2 months to 1 year.

The dust in the atmosphere would have blocked the penetration of solar radiation and led

to extreme cooling in continental areas, with temperatures below freezing in locales distant from the oceans. The oceans themselves possess enough heat capacity to prevent their temperatures falling by more than a few degrees. Other studies suggest that, after the cooling, water vapor lofted by the impact (especially if it was oceanic) might have led to an enhanced greenhouse effect that then raised surface temperatures to the point where the remaining life was severely stressed again.

Computations of the amount of NO produced by the impact (using methods discussed in the last section) have suggested to other workers that the NO<sub>x</sub> produced by the asteroid's shock wave might have combined with water vapor to generate large quantities of nitric acid, HNO<sub>3</sub>. This nitric acid would precipitate onto the surface in the form of very acid rain with pH ranging from 0 to about 1.5. Such acid conditions could well have exterminated marine organisms with calcareous shells in addition to damaging the leaves of vegetation and the lungs of animals. Finally, the evidence for global wildfires and massive injections of soot into the atmosphere also indicates conditions inimical to life and the possibility of yet other killing mechanisms.

The havoc recorded in the Cretaceous-Tertiary boundary clays is probably not unique in earth history. Estimates of the terrestrial cratering rate (Sec. 10.2.2) suggest that a 10-km diameter or larger object impacts the earth with a frequency of about once per 100 Myr. No unequivocal evidence has yet been found that other great biological extinctions were caused by impacts, but marine cores have provided evidence for some local extinctions caused by smaller events. The recognition that large impacts may have played an important role in the history of life on earth has spurred geologists to search more diligently for evidence of such events, and new and exciting results are to be expected in the coming years.

The study of the effects of a large impact on the earth also led to the realization that a full-scale nuclear war might produce the same types of climatic disturbances. Studies of the effects of "nuclear winter" (Turco et al., 1983) show that they might be comparable to those produced by the Cretaceous-Tertiary impact and lead any intelligent person to the conclusion that no nation can afford to fight such a war.

## 12.11 THE GIANT IMPACT THEORY OF THE MOON'S ORIGIN

A final development that underscores the importance of the impact cratering process in solar system history is the recent proposal that the earth's moon formed as the result of an impact between the protoearth and a Mars-size protoplanet. A sce-

## IMPACT CRATERING

nario of this kind would have been unthinkable a few years ago, but studies of planetary accretion discussed in Section 12.2 have shown the likelihood that the later stages of accretion are punctuated by very large impacts between objects of planetary dimensions.

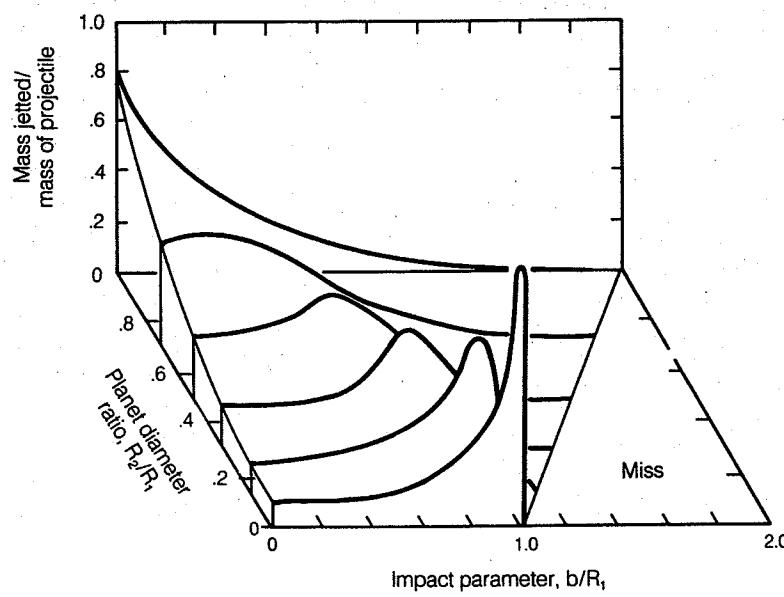
The first suggestions that the moon may have originated in a giant impact were made independently by Hartmann and Davis (1975) and Cameron and Ward (1976). These suggestions were prompted by the analysis of samples brought back from the moon's surface by the 1969–1972 Apollo missions. Defying predictions made by the three classic theories of the moon's origin (fission, capture, and coaccretion), these samples showed that the chemistry of the lunar rocks is basically similar to the earth's mantle but with notable enrichments in refractory elements and depletions of volatile elements. The volatile depletions reach an extreme in the case of water, where no evidence for water of any kind, not even water of crystallization, has been found in the lunar samples. These facts, coupled with the moon's lack of a significant iron core, are inexplicable on the basis of any of the three classic models.

The peculiar chemistry of the lunar rocks had earlier prompted Ringwood (1970, 1972, 1984) to suggest that the moon was composed of mantle material that had been somehow "boiled off" the earth and condensed in space, a process that would have depleted the volatiles and enriched the refractory elements. Ringwood's proposals were not taken seriously by most planetary sci-

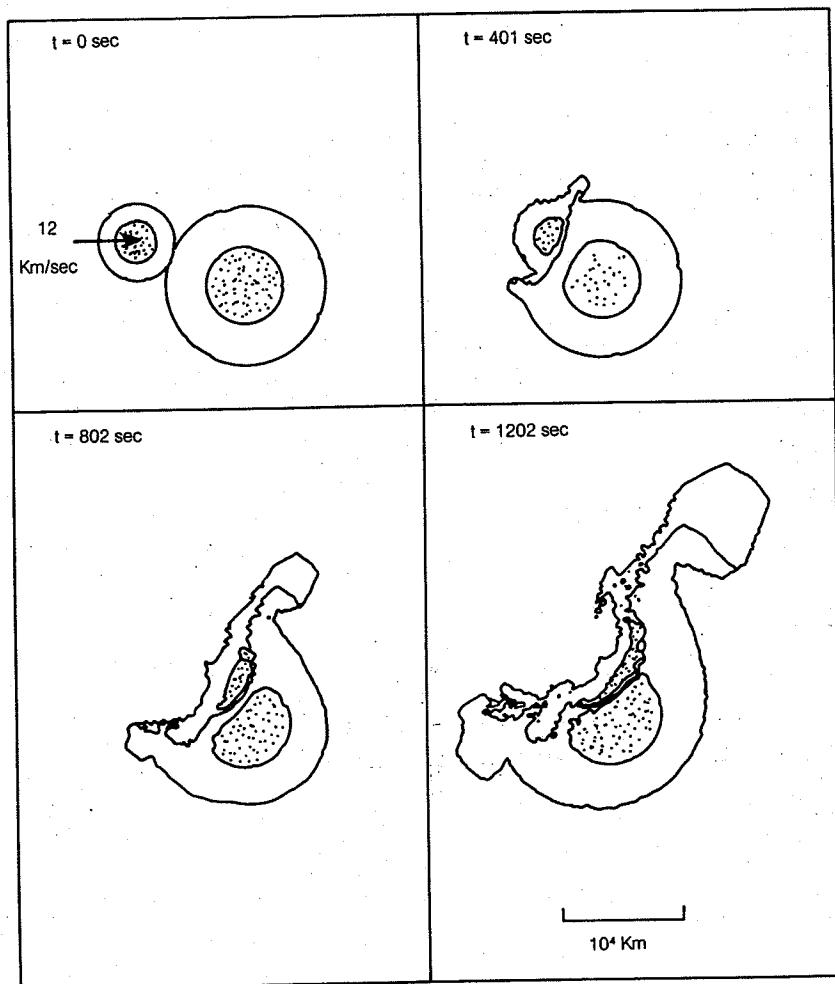
entists because his dynamical mechanism for injecting this material into orbit was faulty.

Once such an event became thinkable, however, it was soon realized that a single giant impact might be capable of lofting enough material into earth orbit to form the moon. In this process it is not the conventional ejecta curtain that finally makes it into orbit, because its maximum ejection velocity is only about 10 percent of the impact velocity, which is slower than the minimum orbital velocity. Nor can high-speed spalled ejecta fragments achieve orbit, because as solid particles they follow Keplerian orbits that are either closed ellipses that return to their starting point (i.e., crash on the surface) or escape the earth entirely. The only material that has a good chance of making it into orbit is vapor ejected from the interface between the projectile and target planets by the process of jetting (Sec. 4.4). Because this vapor expands under the influence of internal pressure gradients in addition to gravity and its own inertia, parcels of the vapor do not follow Keplerian orbits until after they condense into dust. Portions of this high-speed vapor cloud may thus achieve stable orbits above the Roche limit and eventually condense into an orbiting satellite.

The jetted vapor originates from the interface between the protoearth and the projectile, and so it cannot contain metallic iron if the earth's core has already formed. As the vapor plume condenses in space, volatile elements, the last to condense, disperse much farther than silicate vapor



**Fig. 12.5** Total mass jetticed from the interface between two spherical planets as a function of the ratio of the planet's radii,  $R_2/R_1$ , where planet 1 is the larger of the two, and the impact parameter  $b$  normalized by the radius of planet 1. The impact parameter is the offset between the projectile's projected line of travel and the target planet's center. The jetticed mass is a strongly peaked function of impact parameter. After Melosh and Sonett (1986).



**Fig. 12.6** Four stages in the collision of a Mars-size protoplanet with the protoearth at an impact velocity of 12 km/second and impact parameter equal to 0.59 of the earth's radius. Each frame is separated by about 400 seconds from the previous one. A hot jet of vaporized mantle material forms during the collision and expands away from the impact site. This jet contains material from both the projectile and target planets' mantles, but no core material. *From 1987 unpublished computations by M. E. Kipp and H. J. Melosh.*

and are liable to be swept up by the solar wind and thus depleted in the dust ring that forms around the earth. Refractory elements are conversely enriched. Once the dust ring forms it is subject to the same kind of accumulation processes as planetesimals in the early solar system, with the difference that the timescale is much shorter because of the relatively short orbital period of debris in earth orbit. Accumulation into a moon-size body may thus have occurred over a period of only tens to hundreds of years after the putative giant impact event.

Figure 12.5 is the result of a semianalytic computation of the amount of vapor ejected from an impact between two protoplanets as a function of impact parameter had diameter ratio. It shows that a total mass approximating half the projectile's mass may be ejected in the jet. Of this material, approximately half originates from the pro-

jectile and half is from the target planet's mantle. Although such computations are instructive, they are based on a variety of approximations that are difficult to justify fully. The only other approach to the study of planetary collisions (direct experiments seems to be out of the question in the foreseeable future!) is to use detailed hydrodynamic numerical codes. Figure 12.6 shows the outcome of one such computation. It is clear from these figures that a vapor jet of the predicted magnitude and composition does form in a planetary collision. The ultimate fate of this plume has not yet been studied in detail.

Numerical computations by other authors (Benz, Slattery, and Cameron, 1986) emphasize the importance of higher-harmonic gravitational forces between the distorted projectile and target after the collision. In these computations the nascent moon is composed mostly of projectile target

material and the apparent close chemical similarities of the moon to the earth's mantle are fortuitous.

The scientific evaluation of the giant-impact hypothesis of the moon's origin is only just beginning. It has had sufficient success already, however, that it is treated as a very serious contender for the honor of being considered the best explanation of how the moon came to be. Throughout the evaluation of this hypothesis, a full understanding of impact processes will be imperative. It is likely that a deeper understanding of the physics of impact will also be gained in this endeavor.

#### FURTHER READING

The interesting story of oil production from impact craters is told in

R. R. Donofrio, Impact craters: Implications for basement hydrocarbon production, *J. Petrol. Geol.* 3: 279-302 (1981).

The impact origin of the Sudbury structure, originally proposed by R. S. Dietz in 1965, is described in

W. V. Peredery and G. G. Morrison, Discussion of the Origin of the Sudbury Structure, Chapter 22 of E. G. Pye, A. J. Naldrett, and P. E. Giblin (Eds.), *The Geology and Ore Deposits of the Sudbury Structure*, Ontario Geological Survey Special Vol. 1, Government of Ontario, Ontario, Canada (1984).

The study of planetary accretion has a large literature. A fine overview of the subject is given by G. W. Wetherill, Formation of the terrestrial planets, *Ann. Rev. Astron. Astrophys.* 18: 77-113 (1980).

Wetherill's review does not treat the rotational states of the resulting planets. A review that does an especially good job on this topic is

A. W. Harris and W. R. Ward, Dynamical constraints on the formation and evolution of planetary bodies, *Ann. Rev. Earth Planet. Sci.* 10: 61-108 (1982).

The importance of low-velocity cratering for planetary accretion is well brought out by

W. K. Hartman, Planet formation: Mechanism of early growth, *Icarus* 33: 50-61 (1978).

A frequently cited model for the early stages of planetesimal growth is most clearly expounded in

R. Greenberg, J. F. Wacker, W. K. Hartmann and C. R. Chapman, Planetesimals to planets: Numerical simulation of collisional evolution, *Icarus* 35: 1-26 (1978).

The new appreciation of the importance of giant collisions in planetary accretion has not yet resulted in a large literature. The seminal paper in this field is

G. W. Wetherill, Occurrence of giant impacts during the growth of the terrestrial planets, *Science* 228: 877-879 (1985).

The thermal state of planets growing by moderate-sized impacts is considered in detail for the earth and moon by

W. M. Kaula, Thermal evolution of earth and moon growing by planetesimal impacts, *J. Geophys. Res.* 84: 999-1008 (1979).

Impact erosion of bodies from asteroid size up to planetary scale is well treated in a short article

J. D. O'Keefe and T. J. Ahrens, Meteorite impact ejecta: Dependence of mass and energy lost on planetary escape velocity, *Science* 198: 1249-1251 (1977).

The best overall review of asteroid fragmentation available at the present time is somewhat flawed (the units in Equation 2 are incorrect), but is otherwise useful:

D. R. Davis, C. R. Chapman, R. Greenberg, S. J. Weidenschilling, Collisional evolution of asteroids: Populations, rotations, and velocities, in T. Gehrels (Ed.), *Asteroids*, University of Arizona Press, Tucson, pp. 528-557 (1979).

There has recently been a great surge of interest in the experimental study of small-body fragmentation. The classic paper in this field can still be read with profit:

D. E. Gault and J. A. Wedekind, The destruction of tektites by micrometeroroid impact, *J. Geophys. Res.* 74: 6780-6794 (1969).

Several Japanese groups have made important contributions in this field. One good example is

T. Matsui, T. Waza, K. Kani, and S. Suzuki, Laboratory simulation of planetesimal collisions, *J. Geophys. Res.* 87: 10968-10982 (1982).

The impact-induced seismic modification of planetary surface features is discussed by

P. H. Schultz and D. E. Gault, Seismically induced modification of lunar surface features, *Proc. Lunar Sci. Conf. 6th*, pp. 2845-2862 (1975).

The effect of very large impacts at the antipodes of the planet on which they form is addressed by the same two authors in

P. H. Schultz and D. E. Gault, Seismic effects from major basin formations on the moon and Mercury, *The Moon* 12: 159-177 (1975).

Study of the effects of impacts on the origin of atmospheres is still in its infancy. An important contribution to this field was made by

M. A. Lange and T. J. Ahrens, The evolution of an impact-generated atmosphere, *Icarus* 51: 96-120 (1982).

The effects of a large impact on the earth have received a great deal of attention since the original paper by Alvarez et al. (1980). A great many papers on this subject can be found in the proceedings of a meeting convened in Snowbird, Utah, shortly after the discovery by Alvarez et al.

L. T. Silver and P. H. Schultz (Eds.), *Geological implications of impacts of large asteroids*

*and comets on the earth, Geol. Soc. Amer. Special Paper 190 (1982).*

A recent assessment of this theory with copious references to the literature (including the nuclear winter idea) is

W. Alvarez, Toward a theory of impact crises, *Eos* 67: 649, 653-655, 658 (1986).

A relatively complete numerical computation of the effects of a 10-km diameter asteroid striking the earth vertically at 20 km/second has been published by

D. J. Roddy, S. H. Schuster, M. Rosenblatt, L. B. Grant, P. J. Hassig, and K. N. Kreyen-hagen, Computer simulations of large asteroid impacts into oceanic and continental sites—Preliminary results on atmospheric,

cratering and ejecta dynamics, *Int. J. Impact Engng.* 5: 525-541 (1987).

Another computational study of a large impact in an ocean that emphasizes wave formation is

T. J. Ahrens and J. D. O'Keefe, Impact of an asteroid or comet in the ocean and extinction of terrestrial life, *J. Geophys. Res.* 88, Suppl.: A799-A806 (1983).

The giant-impact theory of the moon's origin is too new to have much of a literature. The interested reader will find a number of discussions of this idea from different points of view in Section 7 of

W. K. Hartmann, R. J. Phillips, and G. J. Taylor (Eds.), *Origin of the Moon*, Lunar and Planetary Institute, Houston (1986).

## APPENDIX I

### AI DERIVATION OF THE HUGONIOT EQUATIONS

The Hugoniot equations relate the pressure  $P$ , internal energy per unit mass  $E$ , and density  $\rho$  in front of a shock wave ( $P_0$ ,  $E_0$ ,  $\rho_0$ ) to the values of the same variables ( $P$ ,  $E$ ,  $\rho$ ) after the shock wave has passed. The density is sometimes expressed in terms of specific volume  $V = 1/\rho$ , and  $V_0 = 1/\rho_0$ . The initial pressure, internal energy, and density are assumed to be known. In addition to the final pressure, energy, and density, the shock velocity  $U$  and particle velocity  $u_p$  behind the shock are unknown (the reference frame is usually chosen so that the unshocked material is at rest). The Hugoniot equations use the conservation of mass, momentum, and energy across the shock front to reduce the number of unknowns from five to two. The equation of state then provides a relation between the pressure, internal energy, and density to completely determine the conditions behind the shock wave.

Many derivations of the Hugoniot equations have been published. I have found the one below to be one of the easiest for students to follow.

#### AI.1 Mass conservation

Figure AI.1 illustrates a block of material through which a shock wave is passing. This is a “free-body” diagram in the sense that all forces acting on the block are explicitly shown. The cross-sectional area of the block  $A$  is constant as the shock moves through it. The pressures  $P$  and  $P_0$  on the block’s sides are not shown in the figure to avoid clutter: they are completely balanced and so play no direct role in driving the shock, so they are ignored in this derivation. Only the pressures acting on the ends of the block, in the direction of the shock wave’s motion, are significant.

The figure shows the block at two different times,  $t$  and  $t'$ . At the earliest time  $t$ , the length of the unshocked region is  $l_u$  and the length of the shocked region is  $l_s$ . Later, at time  $t'$ , the shock wave has progressed a distance  $U(t' - t)$  farther to the right and the shocked end of the block, moving at the particle velocity  $u_p$ , has progressed  $u_p(t' - t)$  farther to the right. The unshocked end of the block, assumed to be at rest, has not moved. The new lengths of the unshocked region

$l'_u$  and shocked region  $l'_s$  are thus given by

$$l'_u = l_u - U(t' - t) \quad (\text{AI.1.1a})$$

$$l'_s = l_s + U(t' - t) - u_p(t' - t) \quad (\text{AI.1.1b})$$

The mass contained in the unshocked portion of the block at time  $t$  is its volume  $l_u A$  times its density  $\rho_0$ ,  $\rho_0 l_u A$ . The mass in the shocked portion is likewise  $\rho l_s A$ . Mass conservation just means that the masses at times  $t$  and  $t'$  must be equal. That is,

$$\rho l_u A + \rho_0 l_u A = \rho l'_s A + \rho_0 l'_u A \quad (\text{AI.1.2})$$

Cancelling through the common factor  $A$ , substituting Equations AI.1.1a and AI.1.1b for  $l'_s$  and  $l'_u$  in Equation AI.1.2, then canceling the terms  $\rho l_s$  and  $\rho_0 l_u$  on both the left and right hand sides,

$$0 = \rho (U - u_p) (t' - t) - \rho_0 U (t' - t) \quad (\text{AI.1.3})$$

Finally, canceling the common factor  $(t' - t)$  and rearranging, we obtain the first Hugoniot equation (3.4.1) of the text

$$\rho (U - u_p) = \rho_0 U \quad (\text{AI.1.4})$$

#### AI.2 Momentum conservation

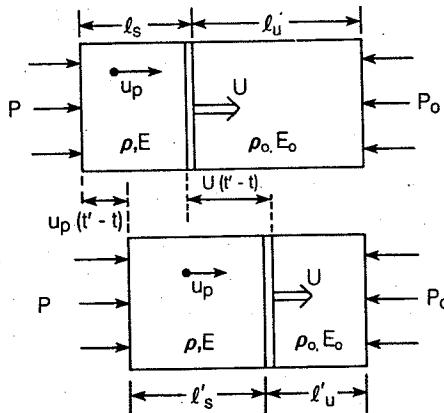
Pressure  $P$  on the shocked end of the block in Figure AI.1 is larger than the pressure  $P_0$  on the unshocked end, so a net force  $F = (P - P_0) A$  acts toward the right, accelerating material in that direction. The momentum of material in the block at time  $t$ ,  $\rho l_s u_p A$ , is thus not equal to the momentum  $\rho l'_s u_p A$  at time  $t'$ . The difference is equal to the momentum imparted by the applied force  $F$  over the time interval  $t' - t$ ,  $F(t' - t)$ . The net momentum balance is thus

$$\rho l'_s u_p A - \rho l_s u_p A = (P - P_0) A(t' - t) \quad (\text{AI.2.1})$$

Cancelling through the area  $A$  and substituting for  $l'_s$  using Equation AI.1.1b, noting that the term  $\rho l_s u_p$  is subtracted from itself on the left side of the equation, and further canceling the common factor  $(t' - t)$  from the remaining terms we obtain

$$\rho (U - u_p) u_p = (P - P_0) \quad (\text{AI.2.2})$$

Now use Equation AI.1.4 to replace  $\rho(U - u_p)$  by  $\rho_0 U$  and rearrange to obtain the second Hugoniot



**Fig. A1.1** Free body diagrams of a shock wave passing through a mass of material at times  $t$  (above) and  $t'$  (below).

equation (3.4.2) in the text,

$$P - P_0 = \rho_0 U u_p \quad (\text{AI.2.3})$$

### A1.3. Energy conservation

Like momentum, the total energy in the block at time  $t$  is not equal to that at time  $t'$  because the applied forces do work on the system. This work is equal to the force times the distance through which it acts. Since the displacement of the unshocked end of the block is zero, the total energy gained between  $t$  and  $t'$  is thus  $PAu_p(t' - t)$ , equal to the force  $PA$  on the shocked end of the block times the distance  $u_p(t' - t)$  through which it acts.

The total energy  $E_{\text{tot}}(t)$  in the block at time  $t$  is the sum of the internal energies in the shocked and unshocked portions and the kinetic energy in the shocked portion:

$$\begin{aligned} E_{\text{tot}}(t) &= \rho_0 l_u E_0 A + \rho l_s E \\ &\quad + 1/2 \rho l_s u_p^2 A \end{aligned} \quad (\text{AI.3.1a})$$

Similarly, at time  $t'$  the total energy is

$$\begin{aligned} E_{\text{tot}}(t') &= \rho_0 l'_u E_0 A + \rho l'_s E \\ &\quad + 1/2 \rho l'_s u_p^2 A \end{aligned} \quad (\text{A.3.1b})$$

Energy conservation thus requires

$$E_{\text{tot}}(t') - E_{\text{tot}}(t) = PAu_p(t' - t) \quad (\text{A1.3.2})$$

Substituting Equations AI.3.1a and AI.3.1b into Equation AI.3.2, cancel  $A$  through as before, substitute Equation AI.1.1a for  $l'_u$  and Equation AI.1.1b for  $l'_s$ , and simplify. The common factor  $(t' - t)$  may then be canceled to obtain

$$\begin{aligned} &- \rho_0 E_0 U + \rho E(U - u_p) \\ &\quad + 1/2 \rho u_p^2 (U - u_p) = P u_p \end{aligned} \quad (\text{AI.3.3})$$

Now replace  $\rho(U - u_p)$  by  $\rho_0 U$  using the first Hugoniot equation (AI.1.4) to obtain

$$\rho_0 U(E - E_0) + 1/2 \rho_0 u_p^2 U = P u_p \quad ((\text{AI.3.4}))$$

We then proceed using two auxiliary equations that can be readily derived from the first two Hugoniot equations (AI.1.4) and (AI.2.3) by eliminating either  $U$  or  $u_p$ , respectively, between the two equations:

$$u_p = \sqrt{(P - P_0)(V - V_0)} \quad (\text{AI.3.5a})$$

and

$$U = 1/\rho_0 \sqrt{(P - P_0)/(V - V_0)} \quad (\text{AI.3.5b})$$

where  $V = 1/\rho$  and  $V_0 = 1/\rho_0$  are the specific volumes of the shocked and unshocked material, respectively. Substituting Equation AI.3.5a for  $u_p$  and AI.3.5b for  $U$  in Equation AI.3.4, canceling the common factor  $\sqrt{(P - P_0)}$ , and rearranging,

$$E - E_0 = 1/2 (P + P_0)(V_0 - V) \quad (\text{AI.3.6})$$

which is the third, and final, Hugoniot equation (3.4.3) of the text.

## APPENDIX II

### AII EQUATIONS OF STATE FOR IMPACT CRATERING

An equation of state relates the three thermodynamic variables pressure  $P$ , density  $\rho$  (or specific volume  $V = 1/\rho$ ), and specific internal energy  $E$  (or temperature  $T$ ). The equation of state is different for different materials and is a complex function of the molecular and atomic structure of the given substance. The response of a material to the shock of an impact is primarily governed by its equation of state, since the Hugoniot equations derived in Appendix I are the same for all materials.

A great deal of work has been done on the equation of state of materials at high pressure since World War II, primarily because of the needs of nuclear weapon design. The equation of state of many simple materials can now be calculated from solid-state physics with a reasonable degree of accuracy, although the equations of state of complex materials (which include nearly all materials of geologic interest) must still be measured. A number of methods have been evolved to fit the measurements empirically and the fitted parameters have been tabulated so that a reasonable approximation to the equation of state is frequently available. The reader should be cautioned, however, that most of these fits have been established from data obtained at particle velocities less than about 5 km/second, which is at the low end of the range of interest for meteoroid impacts. Although the most sophisticated models use such physically motivated high-pressure approximations as the Thomas-Fermi asymptotic limit, there is still considerable uncertainty in the equation of state of geologic materials in the velocity range at which impacts occur in the solar system.

Numerical hydrocode computations typically yield the material density  $\rho$  and specific internal energy  $E$  as the result of one cycle of computation, from which the pressure  $P$  must be computed in the next cycle. The most convenient form of the equation of state for impact computations is thus

$$P = P(\rho, E).$$

The temperature  $T$  is only indirectly related to the specific internal energy  $E$ . Since temperature is not available directly in hydrocode computations, the form of the equation of state involving  $E$  is preferred. The temperature of shock-compressed

material must be estimated from the internal energy by integration of the equation of state and is generally not considered very reliable. Recent shock wave studies are improving this situation by measuring the temperature directly, which in turn allows the equation of state to be refined.

#### AII.1 The liquid equation of state

The simplest equation of state is that of a liquid subjected to a small compression or expansion. In this case, the pressure is a linear function of the density fluctuation about the uncompressed density  $\rho_0$ :

$$P = c_B^2 (\rho - \rho_0) \quad (\text{AII.1.1})$$

where the constant  $c_B$  is the bulk sound speed,  $c_B = (K_0/\rho_0)^{1/2}$ . Equation of state AII.1.1 is a function of density, but has no dependence on internal energy. This lack of dependence on internal energy is only apparent, however, because the bulk modulus  $K_0$  is a weak function of temperature. For sound waves  $K_0$  is the isentropic bulk modulus, whereas for slow static compression  $K_0$  is the isothermal bulk modulus. These two moduli differ by less than a percent for most substances.

Note that densities  $\rho$  much less than  $\rho_0$  often occur in hydrocode computation during rapid expansion. Physically, such densities indicate that the material in a computational cell is no longer a continuum, but may be in the form of droplets or solid fragments separated by gas or vacuum instead. The pressure computed from Equation AII.1.1 must be limited to avoid artificially large tensions in this case. The normal expedient is to test the ratio  $\rho/\rho_0$ ; if it is less than some fixed amount, often chosen between 0.8 and 0.95, the pressure is set to zero or, in sophisticated equations of state, to the vapor pressure over the liquid or solid. This modification should be considered implicit in all the equations of state (except the perfect gas) that follow.

#### AII.2 The perfect gas

The pressure in a perfect gas vanishes at zero temperature so that, in contrast to the liquid equation of state, it depends strongly on the specific internal energy:

$$P = (\gamma - 1) \rho E \quad (\text{AII.2.1})$$

where  $\gamma$  is the ratio of the specific heat at constant pressure to that at constant volume,  $\gamma = C_p/C_v$ . The constant  $\gamma$  is determined by the number of internal degrees of freedom possessed by the gas molecules. For monatomic gases  $\gamma = 1.67$  (no internal degrees of freedom),  $\gamma = 1.4$  for diatomic gases (two internal degrees of freedom), and  $\gamma = 1.3$  for triatomic or more complex gases (three internal degrees of freedom).

### AII.3 The Mie-Gruneisen equation of state

The liquid equation of state (AII.1.1) and the perfect gas equation of state (AII.2.1) are sometimes combined to yield a more realistic equation of state for solids called the "stiffened gas" equation of state:

$$P = c^2(\rho - \rho_0) + (\gamma - 1)\rho E \quad (\text{AII.3.1})$$

This equation has rudimentary dependences on both density and internal energy.

The stiffened gas equation of state above is a special case of the more general Mie-Gruneisen equation of state. In this widely used formulation, the pressure and internal energy are separated into thermal and nonthermal parts,

$$P = P_T + P_c \quad (\text{AII.3.2a})$$

and

$$E = E_T + E_c \quad (\text{AII.3.2b})$$

Where the  $T$  subscript denotes the thermal part and the  $c$  subscript stands for "cold": that is, nonthermal.

The Mie-Gruneisen equation of state is

$$P = P_c + \Gamma \rho (E - E_c) \quad (\text{AII.3.3})$$

where  $\Gamma = \Gamma(\rho)$  is the Gruneisen parameter. Using equations AII.3.2 the Mie-Gruneisen equation of state can be rewritten  $P_T = \Gamma \rho E_T$ , which is the same form as the perfect gas equation of state where  $\Gamma = \gamma - 1$ . In this equation of state the thermal pressure is thus similar to that of a perfect gas, while there is also a pressure due to bulk compression. Note that at zero compression  $P_c = 0$  so that there is no pressure without compression.

At low pressure the Gruneisen parameter  $\Gamma$  can be shown to be

$$\Gamma = \alpha K_0 / \rho_0 C_v \quad (\text{AII.3.4})$$

where  $\alpha$  is the volume coefficient of expansion.  $\Gamma$  is usually about 2 for most solids and  $\frac{5}{3}$  for monatomic gases.  $\Gamma(\rho)$  is difficult to measure in shock-wave experiments, so the following crude approximation is often used

$$\rho \Gamma(\rho) \approx \rho_0 \Gamma(\rho_0) \quad (\text{AII.3.5})$$

It can be shown that  $\Gamma$  is related to the logarithmic derivative of the mean frequency of vibrations in the solid,  $\langle v \rangle$ , with respect to specific volume

$$\Gamma(V) = -\frac{\partial \ln \langle v \rangle}{\partial \ln V} \quad (\text{AII.3.6})$$

The mean vibration frequency  $\langle v \rangle$  can be calculated from solid state theory. If  $\Gamma(\rho)$  is thus known or calculated, and if the cold pressure  $P_c$  and energy  $E_c$  are likewise known or calculated, then the thermal energy  $E_T = E - E_c$  can be determined from shock wave experiments and the temperature can be computed from  $E_T$  using the specific heat at constant volume  $C_v$ , which itself must be calculated. It is no wonder that shock temperatures are thus difficult to determine accurately.

The Mie-Gruneisen equation of state was very popular among shock wave experimenters and theorists alike until very recently. It has a large literature, the best papers of which are Rice et al. (1958), Al'tschuler (1965), and Zel'dovich and Raizer (1967), Chapter 11. Most of these papers examine its implications for shock wave studies and show how its parameters may be derived from measured Hugoniot curves.

### AII.4 Murnaghan equation of state

The Murnaghan equation of state goes a step beyond the liquid equation of state in adding a nonlinear dependence of pressure on density. It does not possess any overt dependence on specific energy and so is mostly limited to describing the results of static isothermal compression. It has been widely used in geophysical work on planetary interiors, but is sometimes used in shock wave computations where a nonlinear equation of state is needed to permit the development of true shock waves but where thermal pressure is an unwanted complication.

The most useful form of this equation is

$$P = (K_0/n)[(\rho/\rho_0)^n - 1] \quad (\text{AII.4.1})$$

where  $K_0$  is the zero-pressure bulk modulus and  $n$  is a constant. Murnaghan parameters for a number of geologically interesting substances are tabulated in Table AII.1.

### AII.5 The Thomas-Fermi high-pressure limit

As a solid is compressed its nuclei are pushed closer together and its electrons, obeying the dictates of the uncertainty principle, become more energetic. With increasing compression the stage is eventually reached where all of the atoms are ionized and the solid is essentially an electron gas in which the immersed nuclei serve mainly to guarantee charge neutrality. Most of the pressure is due to the electrons. The pressure of a cold solid under these circumstances is easy to calculate and is given by

$$P = \frac{\hbar^2}{5m_e} \left( \frac{3}{8\pi} \right)^{2/3} \left( \frac{\rho Z N_0}{\mu} \right)^{5/3} \quad (\text{AII.5.1})$$

TABLE AII.1 *Murnaghan equation of state parameters*

Material	$\rho_0$ (kg/m <sup>3</sup> )	$K_0$ (GPa)	$n$
Iron*	7680	113.5	5.32
Aluminum*	2750	77.2	4.48
Diabase*	3000	60.2	3.76
Basalt*	2860	19.3	5.5
Serpentinite†	2800	20.9	6.0
Granite*	2630	35.7	3.94
Calcite ("carbonate")*	2670	38.5	4.68
Permafrost (water saturated)*	1960	12.3	4.15
Coconino Sandstone*	2000	4.5	4.72
Dry Sand*	1600	4.6	4.24
Ice (0°C)*	910	1.502	5.23
Water	1000	2.2	5.40

\*Kieffer and Simonds (1980)

†Clark (1966)

where  $h$  is Planck's constant,  $m_e$  is the electron's mass,  $Z$  is the mean atomic number,  $\mu$  is the mean atomic weight, and  $N_0$  is Avogadro's number. Assuming that the atomic weight is roughly twice the atomic number times the mass of a nucleon and evaluating the constants, this pressure is (Landau and Lifshitz, 1980, Chapter 11)

$$P \approx 3.1 \times 10^7 \rho^{5/3} \quad (\text{AII.5.2})$$

where  $P$  is in Pascals and  $\rho$  is in kg/m<sup>3</sup>. This equation is probably only good at pressures exceeding about 10,000 GPa.

The Thomas-Fermi theory (and the Thomas-Fermi-Dirac theory, which corrects the Thomas-Fermi theory for electron exchange forces) can be extended down to about 1000 GPa by treating the interaction between the electrons and nuclei more exactly. Finite temperatures can also be treated. This field has a large literature of its own that we do not need to review here. See the review by Gilvarry (1969). The pressure reached in most meteorite impacts is not high enough for the Thomas-Fermi theory to be useful.

#### AII.6 Linear shock-particle velocity relation

Another formulation of the equation of state is founded directly on the results of shock-wave measurements. It is frequently observed that when the shock-wave velocity  $U$  is plotted against the particle velocity  $u_p$  in a shock-wave experiment a linear relation results

$$U = C + Su_p \quad (\text{AII.6.1})$$

where  $C$  is a constant with dimensions of velocity and  $S$  is a dimensionless constant. This equation can be expressed in terms of pressure, density, and internal energy using the Hugoniot equations (3.4.1) through (3.4.3).

A number of values of  $C$  and  $S$  are listed for materials of geologic interest in Table AII.2. Note that even when the  $U - u_p$  relation is not a straight line because of high-pressure phase transitions it can usually be represented by straight-line segments. Table AII.2 gives a description of this type for Quartz, Bronzitite, and Dunite.

TABLE AII.2 *Linear shock-particle velocity equation of state parameters*

Material	$\rho_0$ (kg/m <sup>3</sup> )	$C$ (km/sec)	$S$	Pressure Range (GPa)
Iron*	7680	3.80	1.58	—
Aluminum*	2750	5.30	1.37	—
Diabase*	3000	4.48	1.19	—
Basalt*	2860	2.60	1.62	—
Serpentinite†	2800	2.73	1.76	—
Granite*	2630	3.68	1.24	—
Calcite ("carbonate")*	2670	3.80	1.42	—
Permafrost (water saturated)*	1960	2.51	1.29	—
Coconino Sandstone*	2000	1.50	1.43	—
Dry Sand*	1600	1.70	1.31	—
Ice (0°C)*	910	1.28	1.56	—
Water*	1000	1.48	1.60	—
Quartz (single crystal)‡	2650	3.68 5.56 1.74	2.12 0.14 1.70	< 15 15–38 > 38
Bronzitite Stillwater, Mont.)‡	3280	6.0 7.3 5.2	1.1 0.2 1.2	< 39.8 39.8–60 > 60
Dunite (Twin Sisters, Wash.)‡	3320	6.6 7.8 4.4	0.9 0.2 1.5	< 44 44–73 > 73

\*These curves are "mean fits" over the full range of pressure, ignoring phase transitions. From Kieffer and Simonds (1980).

†Clark (1966)

‡Kieffer (1977)

The parameters  $C$  and  $S$  in Equation AII.6.1 can be expressed in terms of the parameters of some of the equations of state given previously. Thus,  $C$  is related to the bulk sound speed in both the liquid and Murnaghan equations of state by

$$C = c_B = (K_0/\rho_0)^{1/2} \quad (\text{AII6.2})$$

$S$  is related to the Gruneisen parameter  $\Gamma$  by

$$S = [1 + \Gamma(\rho_0)]/2 \quad (\text{AII6.3})$$

or to the Murnaghan parameter  $n$  by

$$S = (n + 1)/2 \quad (\text{AII6.4})$$

The linear shock-particle velocity relation often provides a good description of the Hugoniot up to 6 km/second, the highest velocities achieved in the laboratory. Because of its simplicity it also lends itself to analytic computation and so is widely used in shock wave work.

### AII.7 The Tillotson equation of state

The final equation of state I will discuss was initially derived by J. H. Tillotson (1962) specifically for high-velocity impact computations. Tillotson designed his equation to duplicate the linear shock-particle velocity relation at low pressures and to extrapolate to the Thomas-Fermi limit at high pressures. This equation of state also has parameters that allow it to describe the unloading of shocked material into the vapor phase.

The Tillotson equation has two different forms, depending upon whether the material is compressed to higher density than its zero-pressure form or expanded to lower density. The form used in the compressed region ( $\rho/\rho_0 \geq 1$ ) and for cold expanded states where the energy density is less than the energy of incipient vaporization,  $E < E_{iv}$ , is

$$P = \left[ a + \frac{b}{(E/(E_0\eta^2) + 1)} \right] \rho E + A\mu + B\mu^2 \quad (\text{AII.7.1})$$

where  $\eta = \rho/\rho_0$ ,  $\mu = \eta - 1$ , and  $a$ ,  $b$ ,  $A$ ,  $B$ , and  $E_0$  are the Tillotson parameters. In spite of the notation,  $E_0$  is not the initial energy density of the substance: it is merely a parameter that is often close to the vaporization energy. The initial energy density  $E$  must actually be zero to assure that  $P = 0$  in the initial state. Note that a low-density pressure cutoff in cold expanded states for  $\rho/\rho_0 < 0.8$  to 0.95 must be applied, as discussed in Sec. AII.1.

It can be shown that  $(a + b)$  is the Gruneisen parameter at zero pressure,

$$(a + b) = \Gamma(\rho_0) \quad (\text{AII.7.2})$$

The parameter  $a$  is usually chosen to equal 0.5 to give the correct high-pressure limit. Note that at high pressure and temperature  $P \rightarrow a\rho E$ , and  $a =$

0.5 is not what would be expected from the perfect gas equation of state (AII.2.1) for a "monatomic" electron gas:  $a = \frac{1}{3}$  is the limit predicted by the Thomas-Fermi model. However,  $a = 0.5$  has been found to fit observational data better and so is conventionally used.

The  $C$  and  $S$  parameters of the linear shock-particle velocity model can be expressed in terms of the Tillotson parameters by

$$C = (A/\rho_0)^{1/2} \quad (\text{AII.7.3})$$

( $A$  is thus equal to  $K_0$ ) and

$$S = \frac{1}{2} [1 + B/A + (a + b)/2] \quad (\text{AII.7.4})$$

In the expanded state,  $\rho/\rho_0 \leq 1$ , when the internal energy exceeds the energy of complete vaporization  $E > E_{cv}$ , the pressure is given by the following equation

$$P = a\rho E + \left\{ \frac{b\rho E}{(E/(E_0\eta^2) + 1)} + A\mu e^{-\beta(\rho_0/\rho - 1)} \right\} e^{-\alpha(\rho_0/\rho - 1)^2} \quad (\text{AII.7.5})$$

where  $\alpha$  and  $\beta$  are constants that control the rate of convergence of this equation to the perfect gas law. Equation AII.7.5 is constructed so that the pressure and its first derivative are continuous with Equation AII.7.1 when  $\rho/\rho_0 = 1$ . However, some trouble may be experienced in the partial vaporization regime when  $\rho/\rho_0 < 1$  and  $E_{iv} < E < E_{cv}$ . In order to make this transition smooth it has been found that in this region the pressure is best computed from a hybrid formula:

$$P = [(E - E_{iv})P_E + (E_{cv} - E_{iv})P_C] / (E_{cv} - E_{iv}) \quad (\text{AII.7.6})$$

where  $P_E$  is computed from the expanded Tillotson equation (AII.7.5) and  $P_C$  is computed from the compressed Tillotson equation (AII.7.1).

Table AII.3 lists the Tillotson parameters for a variety of materials of geologic interest. Also listed are the transformation energies so that pressure can be computed over the full range of density ratios in both the solid and vapor states.

When high-pressure phase transformations occur in a shocked solid the internal energy  $E$  computed from the Hugoniot equation (3.4.3) includes the transition energy  $E_{tr}$  from the low- to high-pressure phases. The internal energy  $E_{hpp}$  in the Tillotson equations for the high-pressure phase must be corrected for this transition energy by using the following relation to compute the internal energy

$$E_{hpp} = P(V_{00} - V)/2 - E_{tr} \quad (\text{AII.7.7})$$

where  $V_{00}$  is the specific volume of the low-pressure phase. The specific energy of the low-pressure phase is taken to be zero at standard temperature and pressure.

TABLE AII.3 *Tillotson equation of state parameters*

Material	$\rho_0$ (kg/m <sup>3</sup> )	a	b	A (GPa)	B (GPa)	$E_0$ (MJ/kg)	$\alpha$	$\beta$	$E_{iv}$ (MJ/kg)	$E_{cv}$ (MJ/kg)
Iron*	7800	0.5	1.5	128	105	9.5	5	5	2.4	8.67
Aluminum*	2700	0.5	1.63	75.2	65	5	5	5	3.0	13.9††
Granite†	2680	0.5	1.3	18	18	16	5	5	3.5	18
Gabbroic Anorthosite‡										
(1pp)	2940	0.5	1.5	71	75	487	5	5	4.72	18.2
(hpp)	3970	0.5	1.3	240	130	1800	5	5	3.19	16.8
Andesite†	2700	0.5	1.3	18	18	16	5	5	3.5	18
Wet Tuff†	1970	0.5	1.3	10	6	11	5	5	3.2	16
Dry Tuff†	1700	0.5	1.3	4.5	3	6	5	5	3.5	18
Limestone†	2700	0.5	0.6	40	67	10	5	5	2.5	14
Halite†	2160	0.5	0.6	25	30	5	5	5	2	15
Alluvium†	2600	0.5	0.8	30	10	6	5	5	3.5	18
Water (0°C)**	998	0.7	0.15	2.18	13.25	7	10	5	0.419	2.69
Ice** (hpp, -10°C)	1293	0.3	0.1	10.7	65	10	10	5	0.773	3.04

\*Tillotson (1962)

†Allen (1967)

‡O'Keefe and Ahrens (1982a)

\*\*O'Keefe and Ahrens (1982b)

††from JANAF tables

### AII.8 Semianalytic equations of state

The modern trend in equations of state is toward the use of increasingly complex computer codes to generate an equation of state that relies on different physical approximations in different domains of validity. One of the best such equations of state is the ANEOS code (Thompson and Lauzon, 1972). One major advantage possessed by this code over the Tillotson equation of state is that in ANEOS the pressure, temperature, and density are derived from the Gibbs free energy, and so are guaranteed to be thermodynamically consistent. Pressures computed by the Tillotson equation, on the other hand, may not satisfy thermodynamic consistency requirements, especially

when phase changes occur. The Tillotson equation cannot correctly represent the pressure and density changes in a two-phase region, as when gas and liquid coexist. ANEOS, on the other hand, does a good job in such a region.

Unfortunately, ANEOS does not have a simple analytic form, so to use it the reader must obtain a copy of the code (it consists of about 3000 lines of FORTRAN code). Be warned, however, that this code has a large number of input parameters, so that establishing the equation of state for a new material can be a formidable task. On the benefit side, this equation of state provides a realistic representation of the behavior of a given substance over a wide range of conditions.

## GENERAL BIBLIOGRAPHY

- Ahrens, T. J. and O'Keefe, J. D. (1977). Equations of state and impact-induced shock-wave attenuation on the moon. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, R. B. Merrill) pp. 639-656. Pergamon Press, New York.
- Ahrens, T. J. and O'Keefe, J. D. (1978). Energy and mass distributions of impact ejecta blankets on the moon and Mercury. *Proc. Lunar Planet. Sci. Conf.* 9th pp. 3787-3802.
- Ahrens, T. J., O'Keefe, J. D., and Gibbons, R. V. (1973). Shock compression of a recrystallized anorthositic rock from Apollo 15. *Proc. Lunar Sci. Conf.* 4th, pp. 2575-2590.
- Allen, R. T. (1967). Equations of state of rocks and minerals. Interim report to DASA under contract DA 49-146-XZ-462. General Dynamics Report #GA MD-7834.
- Al'tshuler, L. V. (1965). Use of shock waves in high-pressure physics. *Soviet Physics US-PEKHI* 8: 52-89.
- Andrews, R. J. (1977). Characteristics of debris from small-scale cratering experiments. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, R. B. Merrill) pp. 1089-1100. Pergamon Press, New York.
- Arvidson, R., Drozd, R. J., Hohenberg, C. M., Morgan, C. J., and Poupeau, G. (1975). Horizontal transport of the regolith, modification of features, and erosion rates on the lunar surface. *The Moon* 13:67-79.
- Austin, M. G., Thomsen, J. M., Ruhl, S. F., Orphal, D. L., Borden, W. F., Larson, S. A., and Schultz, P. H. (1981). Z-model analysis of impact cratering: An overview. In *Multi-ring basins* (eds. P. H. Schultz and R. B. Merrill). *Proc. Lunar Planet. Sci. Conf.* 12A, pp. 197-205.
- Baldwin, R. B. (1963). *The Measure of the Moon*. University of Chicago, Press, Chicago.
- Baldwin, R. B. (1972). The tsunami model of the origin of ring structures concentric with large lunar craters. *Phys. Earth and Planet. Interiors* 6:327-339.
- Baldwin, R. B. (1981). On the tsunami theory of the origin of multi-ring basins. In *Multi-ring basins* (eds. P. H. Schultz and R. B. Merrill). *Proc. Lunar Planet. Sci. Conf.* 12A, pp. 275-288.
- Basaltic Volcanism Study Project (1981). *Basaltic Volcanism on the Terrestrial Planets*. Pergamon Press, New York.
- Benz, W., Slattery, W. L. and Cameron, A. G. W. (1986). The origin of the moon and the single-impact hypothesis. *Icarus* 66:515-535.
- Bjork, R. J. (1961). Analysis of the formation of Meteor Crater, Arizona: A preliminary report. *J. Geophys. Res.* 66:3379-3387.
- Bjork, R. L., Kreyenhagen, K. N., and Wagner, M. H. (1967). Analytical Study of Impact Effects as Applied to the Meteoroid Hazard. NASA Report CR-757, 186 pp.
- Boon, J. D. and Albritton, C. C., Jr. (1937). Meteorite scars in ancient rocks. *Field and Lab.* 5:53-64.
- Borg, I. Y. (1972). Some shock effects in granodiorite to 270 kilobars at the Pile Driver site. In *Flow and fracture of rocks*, AGU Monograph 16, pp. 293-311.
- Bratt, S. R., Solomon, S. C., Head, J. W., and Thurber, C. H. (1985). The deep structure of lunar basins: Implications for basin formation and modification. *J. Geophys. Res.* 90:3049-3064.
- Brennan, C. (1974). Isostatic recovery and the strain rate dependent viscosity in the earth's mantle. *J. Geophys. Res.* 79:3993-4001.
- Bryan, J. B., Burton, D. E., Cunningham, M. E., and Lettis, L. A., Jr. (1978). A two-dimensional computer simulation of hypervelocity impact cratering: Some preliminary results for Meteor Crater, Arizona. *Proc. Lunar Planet. Sci. Conf.* 9th, pp. 3931-3964.
- Cameron, A. G. W. and Ward, W. R. (1976). Origin of the Moon (abs). *Lunar and Planet. Sci. VII*, pp. 120-122.
- Chapman, C. R. and McKinnon, W. B. (1986). Cratering of Planetary Satellites. In *Satellites* (eds. J. A. Burns and M. S. Matthews) pp. 492-580. University of Arizona Press, Tucson.
- Charters, A. C. and Summers, J. L. (1959). Comments on phenomena of high-speed impact. U.S. Naval Ordinance Laboratory Report NOLR-1238, pp. 200-221.
- Clark, S. P. (1966). Handbook of Physical Constants (Revised edition). *Geol. Soc. Amer. Memoir* 97.
- Croft, S. K. (1981). The modification stage of basin formation: Conditions of ring formation. In *Multi-ring basins* (eds. P. H. Schultz and R. B. Merrill). *Proc. Lunar Planet. Sci. Conf.* 12A, pp. 227-257.

## IMPACT CRATERING

- Duvall, G. E. and Graham, R.A. (1977). Phase transitions under shock-wave loading. *Rev. Mod. Phys.* 49:523-579.
- Dvorak, J. J. (1979). Analysis of small scale lunar gravity anomalies: Implications for crater formation and crustal history. Caltech PhD Thesis.
- Fink, J. H., Greeley, R., and Gault, D. E. (1981). Impact cratering experiments in bingham materials and the morphology of craters on Mars and Ganymede. *Proc. Lunar Planet. Sci.* 12B:1649-1666.
- Frankel, A. and Clayton, R. W. (1986). Finite difference simulations for seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity. *J. Geophys. Res.* 91:6465-6489.
- Gaffney, E. S. and Melosh, H. J. (1982). Noise and target strength degradation accompanying shallow-buried explosions. *J. Geophys. Res.* 87:1871-1879.
- Gault, D. E. (1970). Saturation and equilibrium conditions for impact cratering on the lunar surface: Criteria and implications. *Radio Science* 5:273-291.
- Gault, D. E. (1974). Impact Cratering. In *A primer in lunar geology* (eds. R. Greeley and P. H. Schultz) pp. 137-175. NASA Ames, Moffett Field.
- Gault, D. E., and Greeley, R. (1978). Exploratory experiments of impact craters formed in viscous-liquid targets: Analog for Martian Rama-part Craters? *Icarus* 34:486-495.
- Gault, D. E., Guest, J. E., Murray, J. B., Dzurisin, D., and Malin, M. C. (1975). Some comparisons of impact craters on Mercury and the Moon. *J. Geophys. Res.* 80:2444-2460.
- Gault, D. E., Hötz, F., Brownlee, D. E., and Hartung, J. B. (1974). Mixing of the lunar regolith. *Proc. Lunar Planet. Sci. Conf.* 5th, pp. 2365-2386.
- Gault, D. E., Shoemaker, E. M., and Moore, H. J. (1963). Spray ejected from the lunar surface by meteoroid impact. NASA Tech. Note D-1767.
- Gault, D. E. and Wedekind, J. A. (1978). Experimental studies of oblique impact. *Proc. Lunar Planet. Sci. Conf.* 9th, pp. 3843-3875.
- Gault, D. E. and Wedekind, J. A. (1979). Experimental results for effects of gravity on impact crater morphology (Abs), *Second Int. Colloquium on Mars*, NASA conference pub 2072, p. 29.
- Gilbert, G. K. (1893). The moon's face: A study of the origin of its features. *Bull. Phil. Soc. Wash.* 12:241-292.
- Gilvarry, J. J. (1969). Equations of state at high pressure from the Thomas-Fermi atom model. In *The application of modern physics to the earth and planetary interiors* (ed. S. K. Runcorn) pp. 313-403. Wiley-Interscience, London.
- Golombeck, M. P. (1985). Fault type predictions from stress distributions on planetary surfaces: Importance of initiation depth. *J. Geophys. Res.* 90:3065-3074.
- Grady, D. E. (1977). Processes occurring in shock wave compression of rocks and minerals. In *High pressure research: Applications in geophysics* (eds. M. H. Managnani and S.-I. Akimoto), pp. 389-438. Academic Press, New York.
- Grady, D. E. (1980). Shock deformation of brittle solids. *J. Geophys. Res.* 85:913-924.
- Grady, D. E. and Murri, W. J. (1976). Dynamic unloading in shock compressed feldspar. *Geophys. Res. Lett.* 3:472-474.
- Greeley, R., Fink, J., Gault, D. E., Snyder, D. B., Guest, J. E., and Schultz, P. (1980). Impact cratering in viscous targets: Laboratory experiments. *Proc. Lunar Planet. Sci. Conf.* 11th, pp. 2075-2097.
- Greenberg, R., Wacker, J. F., Hartmann, W. K., and Chapman, C. R. (1978). Planetesimals to planets: Numerical simulation of collisional evolution. *Icarus* 35:1-26.
- Grieve, R. A. F. (1987). Terrestrial impact structures. *Ann. Rev. Earth Planet. Sci.* 15:245-270.
- Grieve, R. A. F., Dence, M. R., and Robertson, P. B. (1977). Cratering processes: As interpreted from the occurrence of impact melts. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, and R. B. Merrill) pp. 791-814. Pergamon Press, New York.
- Grieve, R. A. F. and Head, J. W. (1983). The Manicouagan impact structure: An analysis of its original dimensions and form. *J. Geophys. Res.* 88, Suppl.:A807-A818.
- Grieve, R. A. F., Robertson, P. B., and Dence, M. R. (1981). Constraints on the formation of ring impact structures, based on terrestrial data. In *Multiring Basins* (eds. P. H. Schultz and R. B. Merrill). *Proc. Lunar Planet. Sci. Conf.* 12A, pp. 37-57.
- Grün, E., Zook, H. A., Fechtig, H., and Giese, R. H. (1985). Collisional balance of the meteoritic complex. *Icarus* 62:244-272.
- Hager, D. (1953). Crater mound (meteor crater), Arizona, a geologic feature. *Bull. of Am. Assoc. Petrol. Geol.* 37, 821-857.
- Hale, W. and Grieve, R. A. F. (1982). Volumetric analysis of complex lunar craters: Implications for basin ring formation. *J. Geophys. Res.* 87, Suppl.:A65-A76.
- Hale, W. and Head, J. W. (1980a). Central peaks in mercurian craters: Comparisons to the moon. *Proc. Lunar Planet. Sci. Conf.* 11th, pp. 2191-2205.
- Hale, W. and Head, J. W. (1980b). Crater central peaks on the moon, Mercury, and earth. Reports on the Planetary Geology Program, NASA Technical Memorandum 82385, pp. 131-133.

- Hale, W. S. (1983). Central structures in Martian impact craters: Morphology, morphometry, and implications for substrate volatile distributions. *Lunar and Planetary Sci. Conf. XIV* (abs), pp. 273-274.
- Hall, J. L., Solomon, S. C., and Head, J. W. (1981). Lunar floor-fractured craters: Evidence for viscous relaxation of crater topography. *J. Geophys. Res.* 86:9537-9552.
- Hartmann, W. K. and Davis, D. R. (1975). Satellite-sized planetesimals and lunar origin. *Icarus* 24:504-515.
- Hartmann, W. K. and Wood, C. A. (1971). Moon: Origin and evolution of multi-ring basins. *The Moon* 3:3-78.
- Hartmann, W. K. and Yale, F. G. (1968). Mare Orientale and its basin system. *Comm. Lunar Planetary Lab* 7:131-137.
- Head, J. W. (1974). Orientale multi-ringed basin interior and implications for the petrogenesis of lunar highland samples. *The Moon* 11: pp. 327-356.
- Head, J. W. (1977). Origin of outer rings in lunar multi-ring basins. Evidence from morphology and ring spacing. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, R. B. Merrill) pp. 563-573. Pergamon Press, New York.
- Hodges, C. A. and Wilhelms, D. E. (1978). Formation of lunar basin rings. *Icarus* 34:294-323.
- Holsapple, K. A. and Schmidt, R. M. (1982). On the scaling of crater dimensions 2. Impact processes, *J. Geophys. Res.* 87:1849-1870.
- Holsapple, K. A. and Schmidt, R. M. (1987). Point source solutions and coupling parameters in cratering mechanics. *J. Geophys. Res.* 92:6350-6376.
- Hörz, F., Ostertag, R., and Rainey, D. A. (1983). Bunte breccia of the Ries: Continuous deposits of large impact craters. *Rev. Geophys. and Space Phys.* 21:1667-1725.
- Howard, K. A. (1973). Avalanche mode of motion: Implications from lunar examples. *Science* 180:1052-1055.
- Howard, K. A. (1974). Fresh lunar impact craters: Review of variations with size, in *Proc. Fifth Lunar Conf.* vol 1, pp. 61-69.
- Howard, K. A., Wilhelms, D. E., and Scott, D. H. (1979). Lunar basin formation and highland stratigraphy, *Rev. Geophys. Space Phys.* 12:309-327.
- Hoyt, W. G. (1987). *Coon Mountain controversies*. University of Arizona Press, Tucson.
- Hsü, K. J. (1975). Catastrophic debris streams (Sturzstroms) generated by rockfalls. *Geol. Soc. Amer. Bull.* 86:129-140.
- Ivanov, B. A. (1986). *Cratering mechanics*. English translation of Russian original, NASA Tech. Memorandum 88477 (N87-15662).
- Jeanloz, R. (1980). Shock effects in olivine and implications for Hugoniot data. *J. Geophys. Res.* 85:3163-3176,
- Kaula, W. M. (1979). Thermal evolution of the earth and moon growing by planetesimal impacts. *J. Geophys. Res.* 84:999-1008.
- Kieffer, S. W. (1977). Impact conditions required for formation of melt by jetting in silicates. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, and R. B. Merrill) pp. 751-769. Pergamon Press, New York.
- Kieffer, S. W. and Simonds, C. H. (1980). The role of volatiles and lithology in the impact cratering process. *Rev. Geophys. and Space Phys.* 18:143-181.
- Kopal, Z. and Carder, R. W. (1979). *Mapping of the Moon*, D. Reidel Pub. Co., Holland, 237 pp.
- Landau, L. D. and Lifshitz, E. M. (1980). *Statistical Physics* (3rd ed.) Part 1. Pergamon Press, New York.
- Larson, D. B. (1977). The relationship of rock properties to explosive energy coupling. University of California Research Labs report UCRL-52204.
- Larson, D. B. (1982). Inelastic wave propagation in sodium chloride. *Bull. Seis. Soc. Amer.* 72:2107-2130.
- Latham, G. V., Ewing, M., Press, F., Sutton, G., Dorman, J., Nakamura, Y., Toksoz, N., Lammelein, D., and Duennebier, F. (1972). Passive seismic experiment. In Apollo 16 Preliminary Science Report, NASA SP-315, pp. 9-1 to 9-29.
- Latham, G. V., McDonald, W. G., and Moore, H. J. (1970). Missile impacts as sources of seismic energy on the moon: *Science* 168:242-245.
- Lowe, D. R. and Byerly, G. R. (1986). Early Archean silicate spherules of probable impact origin. *Geology* 14:83-86.
- McGetchin, T. R., Settle, M. and Head, J. W. (1973). Radial thickness variation in impact crater ejecta: Implications for lunar basin deposits. *Earth Planet. Sci. Lett.* 20, pp. 226-236.
- McKinnon, W. B. (1978). An investigation into the role of plastic failure in crater modification. *Proc. Lunar Planet. Sci. Conf.* 9th, pp. 3965-3973.
- McKinnon, W. B. (1981). Application of ring tectonic theory to Mercury and other solar system bodies. In *Multi-ring basins* (eds. P. H. Schultz and R. B. Merrill). *Proc. Lunar Planet. Sci. Conf.* 12A, pp. 259-273.
- McKinnon, W. B. and Melosh, H. J. (1980). Evolution of planetary lithospheres: Evidence from multi-ringed basins on Ganymede and Callisto. *Icarus* 44:454-471.
- Melosh, H. J. (1976). On the origin of fractures radial to lunar basins. *Proc. Lunar Planet. Sci. Conf.* 7th, pp. 2967-2982.
- Melosh, H. J. (1977). Crater modification by gravity: A mechanical analysis of slumping. In *Impact and explosion cratering* (eds. D. J. Roddy,

## IMPACT CRATERING

- R. O. Pepin and R. B. Merrill) pp. 1245-1260. Pergamon Press, New York.
- Melosh, H. J. (1979). Acoustic fluidization: A new geologic process? *J. Geophys. Res.* 84:7513-7520.
- Melosh, H. J. (1981). Atmospheric breakup of terrestrial impactors. In *Multiring Basins* (Eds. P. H. Schultz and R. B. Merrill) *Proc. Lunar Planet. Sci.* 12A, pp. 29-35.
- Melosh, H. J. (1982a). A schematic model of crater modification by gravity. *J. Geophys. Res.* 87:371-380.
- Melosh, H. J. (1982b). A simple mechanical model for Valhalla Basin, Callisto. *J. Geophys. Res.* 87:1880-1890.
- Melosh, H. J. (1983). Acoustic fluidization. *American Scientist* 71:158-165.
- Melosh, H. J. (1984). Impact ejection, spallation and the origin of meteorites. *Icarus* 59:234-260.
- Melosh, H. J. and Gaffney, E. S. (1983). Acoustic fluidization and the scale dependence of impact crater morphology. *J. Geophys. Res.* 88, Suppl.:A830-A834.
- Melosh, H. J. and McKinnon, W. B. (1978). The mechanics of ringed basin formation. *Geophys. Res. Lett.* 5:985-988.
- Melosh, H. J. and Sonett, C. P. (1986). When worlds collide: Jetted vapor plumes and the moon's origin. In *Origin of the moon* (eds. W. K. Hartmann, R. J. Phillips, and G. J. Taylor) pp. 621-642. Lunar and Planetary Institute, Houston.
- Moore, H. J. (1971). Large blocks around lunar craters. In *Analysis of Apollo 10 photography and visual observations*, pp. 26-27. NASA-SP 232, Washington, D.C.
- Moore, H. J. (1976). Missile impact craters (White Sands missile range, New Mexico) and applications to lunar research. U.S.G.S. Prof. Pap. 812-B, 47 pp.
- Moore, H. J., Hodges, C. A., and Scott, D. H. (1974). Multi-ringed basins—Illustrated by Orientale and associated features. *Proc. Lunar Planet. Sci. Conf.* 5th pp. 71-100.
- Nininger, H. H. (1956). *Arizona's Meteorite Crater*. American Meteorite Laboratory, Denver, Colo., 232 pp.
- Nordyke, M. D. (1962). An analysis of cratering data from desert alluvium. *J. Geophys. Res.* 67:1965-1974.
- Oberbeck, V. R. (1975). The role of ballistic erosion and sedimentation in lunar stratigraphy. *Rev. Geophys. Space Phys.* 13:337-362.
- Oberbeck, V. R. and Quaide, W. L. (1968). Genetic implications of lunar regolith thickness variations. *Icarus* 9, 446-465.
- Oberbeck, V. R., Quaide, W. L., Mahan, M. and Paulson, J. (1973). Monte Carlo calculations of lunar regolith thickness distributions. *Icarus* 19:87-107.
- O'Keefe, J. D. and Ahrens, T. J. (1975). Shock effects from a large impact on the moon. *Proc. Lunar Planet. Sci. Conf.* 6th, pp. 2831-2844.
- O'Keefe, J. D. and Ahrens, T. J. (1977). Meteorite impact ejecta: Dependence of mass and energy lost on planetary escape velocity. *Science* 198:1249-1251.
- O'Keefe, J. D. and Ahrens, T. J. (1981). Impact cratering: The effect of crustal strength and planetary gravity. *Rev. Geophys. and Space Phys.* 19:1-12.
- O'Keefe, J. D. and Ahrens, T. J. (1982a). The interaction of the Cretaceous/Tertiary extinction bolide with the atmosphere, ocean, and solid earth. *Geol. Soc. Amer. Special Papers*, 190, pp. 103-120.
- O'Keefe, J. D. and Ahrens, T. J. (1982b). Commentary and meteorite swarm impact on planetary surfaces. *J. Geophys. Res.* 87:6668-6680.
- O'Keefe, J. D. and Ahrens, T. J. (1985). Sampling of planetary material by oblique impact jet entrainment. (Abs.) *Lunar and Planet. Sci. Conf.* XVI, pp. 629-630.
- Passey, Q. R. (1980). Effects of atmospheric breakup on crater field formation. *Lunar and Planetary Sci.* XI, pp. 863-864.
- Passey, Q. R. (1982). Viscosity structure of the lithospheres of Ganymede, Callisto, and Enceladus, and of the earth's upper mantle. Caltech Ph.D. Thesis.
- Passey, Q. R. and Melosh, H. J. (1980). Effects of atmospheric breakup on crater field formation. *Icarus* 42:211-233.
- Passey, Q. R. and Shoemaker, E. M. (1982). Craters and basins on Ganymede and Callisto: Morphological indicators of crustal evolution. In *The satellites of Jupiter* (ed. D. Morrison), pp. 379-434. University of Arizona Press, Tucson.
- Pearce, S. J. and Melosh, H. J. (1986). Terrace width variations in complex lunar craters. *Geophys. Res. Lett.* 13:1419-1422.
- Perret, W. R. and Bass, R. C. (1975). Free-field ground motion induced by underground explosions. Sandia report. SAND74-0252.
- Piekutowski, A. J. (1977). Cratering mechanisms observed in laboratory-scale high-explosive experiments. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, R. B. Merrill) pp. 67-102. Pergamon Press, New York.
- Pieters, C. M. (1982). Copernicus crater central peak: Lunar mountain of unique composition. *Science* 215:59-61.
- Pike, R. J. (1974). Depth/diameter relations of fresh lunar craters: Revision from spacecraft data. *Geophys. Res. Lett.* 1:291-294.
- Pike, R. J. (1976). Crater dimensions from Apollo data and supplemental sources, *The Moon* 15:463-477.
- Pike, R. J. (1977). Size-dependence in the shape of fresh impact craters on the moon. In *Impact*

- and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, and R. B. Merrill) pp. 489-509. Pergamon Press, New York.
- Pike, R. J. (1983). Comment on "A schematic model of crater modification by gravity" by H. J. Melosh. *J. Geophys. Res.* 88:2500-2504.
- Pike, R. J. (1985). Some morphologic systematics of complex impact structures. *Meteoritics* 20:49-68.
- Pike, R. J. (1988). Geomorphology of Impact Craters on Mercury. In *Mercury*, pp. 165-273. University of Arizona Press, Tucson.
- Pohl, J., Stöffler, D., Gall, H. and Ernstson, K. (1977). The Ries impact crater. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, R. B. Merrill) pp. 343-404. Pergamon Press, New York.
- Quaide, W. L. and Oberbeck, V. R. (1968). Thickness determinations of the lunar surface layer from lunar impact craters. *J. Geophys. Res.* 73:5247-5270.
- Raikes, S. A. and Ahrens, R. J. (1979). Post-shock temperatures in minerals. *Geophys. J. Roy. Astron. Soc.* 58:717-747.
- Raizer, Y. P. (1960). Condensation of a cloud of vaporized matter expanding in a vacuum. *Soviet Physics - JETP* 37, pp. 1229-1235.
- Remsberg, A. R. (1982). Tectonics of Valhalla Basin, Callisto. MS Thesis, State University of New York at Stony Brook.
- Rice, M. H., McQueen, R. G. and Walsh, J. M. (1958). Compression of solids by strong shock waves. *Solid State Physics* 10:1-63.
- Ringwood, A. E. (1970). Origin of the moon: The precipitation hypothesis. *Earth Planet. Sci. Lett.* 8:131-140.
- Ringwood, A. E. (1972). Some comparative aspects of lunar origin. *Phys. Earth Plan. Int.* 6:366-376.
- Ringwood, A. E. (1984). Origin of the moon (Abs.). In papers presented to the Conference of the Origin of the Moon, Kona, Hawaii, Oct. 13-16, 1984, p. 46.
- Roddy, D. J. and Davis, L. K. (1977). Shatter cones formed in large-scale experimental explosion craters. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, R. B. Merrill) pp. 715-750. Pergamon Press, New York.
- Schenk, P. M. and McKinnon, W. B. (1987). Ring geometry on Ganymede and Callisto. *Icarus* 72:209-234.
- Schmidt, R. M. (1980). Meteor Crater: Energy of formation—implications of centrifuge scaling. *Proc. Lunar Planet. Sci. Conf.* 11th, pp. 2099-2128.
- Schmidt, R. M. and Housen, K. R. (1987). Some recent advances in the scaling of impact and explosion cratering. *Int. J. Impact Engng.* 5:543-560.
- Schultz, P. H. (1976). Floor-fractured lunar craters. *Moon* 15:241-273.
- Schultz, P. H. and Gault, D. E. (1975). Seismic effects from major basin formation on the moon and Mercury. *The Moon* 12:159-177.
- Schultz, P. H. and Gault, D. E. (1979). Atmospheric effects on martian ejecta emplacement. *J. Geophys. Res.* 84:7669-7687.
- Schultz, P. H. and Gault, D. E. (1985). Clustered impacts: Experiments and implications. *J. Geophys. Res.* 90, 3701-3732.
- Settle, M. and Head, J. W. III (1979). The role of rim slumping in the modification of lunar impact craters. *J. Geophys. Res.* 84:3081-3096.
- Shoemaker, E. M. (1960). Penetration mechanics of high velocity meteorites, illustrated by Meteor Crater, Arizona. In *Rept. of the Int. Geol. Congress, XXI Session, Norden. Part XVIII*, pp. 418-434, Copenhagen.
- Shoemaker, E. M. (1962). Interpretation of lunar craters. In Z. Kopal (Ed.), *Physics and Astronomy of the Moon*, Academic Press, New York and London, pp. 283-359.
- Shoemaker, E. M. (1963). Impact mechanics at Meteor Crater, Arizona. In *The Moon, meteorites, and comets* (eds. B. M. Middlehurst, G. P. Kuiper) pp. 301-336. (The Solar System, vol. 4.) University of Chicago Press, Chicago and London.
- Shoemaker, E. M., Batson, R. M., Holt, H. E., Morris, E. C., Rennilson, J. J., and Whitaker, E. A. (1969). Observations of the lunar regolith and the earth from the television camera on Surveyor 7. *J. Geophys. Res.* 74:6081-6119.
- Shoemaker, E. M., Gault, D. E., and Lunn, R. V., (1961). Shatter cones formed by high speed impact in Dolomite. U.S.G.S. Prof. Pap. 424-D: 365-368.
- Shoemaker, E. M., Hait, M. H., Swann, G. A., Schleicher, D. L., Schaber, G. G., Sutton, R. L., Dahlem, D. H., Goddard, E. N., and Waters, A. C. (1970). Origin of the lunar regolith at Tranquility Base. *Proc. Apollo 11 Lunar Sci. Conf.*, pp. 2399-2412.
- Shreve, R. L. (1966). Sherman Landslide, Alaska. *Science* 154, 1639-1643.
- Soderblom, L. A. (1970). A model for small-impact erosion applied to the lunar surface. *J. Geophys. Res.* 75:2655-2661.
- Solomon, S. C. and Head, J. W. (1980). Lunar mascon basins: Lava filling, tectonics, and evolution of the lithosphere. *Rev. Geophys. and Space Phys.* 18:107-141.
- Stöffler, D. (1972). Deformation and transformation of rock-forming minerals by natural and experimental shock processes. *Fortschritte der Mineralogie* 49:50-113.
- Thompson, S. L. and Lauson, H. S. (1972). Improvements in the Chart D radiation-hydrodynamic CODE III: Revised analytic equations of state. Sandia National Laboratory Report SC-RR-71 0714.
- Tillotson, J. H. (1962). Metallic equations of state

- for hypervelocity impact. General Atomic Report GA-3216.
- Toon, O. B., Pollack, J. B., Ackerman, T. P., Turco, R. P., McKay, C. P., and Liu, M. S. (1982). Evolution of an impact-generated dust cloud and its effects on the atmosphere. *Geol. Soc. Amer. Special Paper* 190, pp. 187-200.
- Turco, R. P., Toon, O. B., Ackerman, T. P., Pollack, J. B., and Sagan, C. (1983). Nuclear winter: Global consequences of multiple nuclear explosions. *Science* 222:1283-1292.
- Turco, R. P., Toon, O. B., Park, C., Whitten, R. C., Pollack, J. B., and Noerdlinger, P. (1982). An analysis of the physical, chemical, optical and historical impacts of the 1908 Tunguska meteor fall. *Icarus* 50:1-52.
- Ullrich, G. W., Roddy, D. J., and Simmons, G. (1977). Numerical simulations of a 20-ton TNT detonation on the earth's surface and implications concerning the mechanics of central uplift formation. In *Impact and explosion cratering* (eds. D. J. Roddy, R. O. Pepin, and R. B. Merrill) pp. 959-982. Pergamon Press, New York.
- Van Dorn, W. G. (1968). Tsunamis on the moon? *Nature* 220:1102-1107.
- Van Dorn, W. G. (1969). Lunar maria: Structure and evolution. *Science* 165:693-695.
- Vickery, A. M. (1986). Effect of an impact-generated gas cloud on the acceleration of solid ejecta. *J. Geophys. Res.* 91:14139-14160.
- Vickery, A. M. (1987). Variation in ejecta size with ejection velocity. *Geophys. Res. Lett.* 14:726-729.
- Wetherill, G. W. (1980). Formation of the terrestrial planets. *Ann. Rev. Astron. Astrophys.* 18:77-113.
- Wetherill, G. W. (1985). Occurrence of giant impacts during the growth of the terrestrial planets. *Science* 228:877-879.
- Whitaker, E. A. (1978). Galileo's lunar observations and the dating of the composition of *Siderius Nuncius*. *J. Hist. Astron.* 9:155-169.
- Wilshire, H. G., and Howard, K. A. (1968). Structural pattern in central uplifts of cryptoexplosion structures as typified by Sierra Madera. *Science* 162:258-261.
- Wood, C. A. (1980). Martian double ring basins: New observations. *Proc. Lunar Planet. Sci. Conf.* 11th, pp. 2191-2205.
- Wood, C. A. (1981). Possible terrestrial analogs of Valhalla and other ripple-ring basins. In *Multi-ring basins* (eds. P. H. Schultz and R. B. Merrill). *Proc. Lunar Planet. Sci. Conf.* 12A, pp. 173-180.
- Wood, C. A. and Head, J. W. (1976). Comparison of impact basins on Mercury, Mars, and the moon. *Proc. Lunar Planet. Sci. Conf.* 7th, pp. 3629-3651.
- Woronow, A. (1977). Crater saturation and equilibrium: A Monte Carlo simulation. *J. Geophys. Res.* 82:2447-2456.
- Zel'dovich, Ya. B. and Raizer, Yu. P. (1967). *The physics of shock waves and high-temperature hydrodynamic phenomena*. Academic Press, New York.

# INDEX

- Ablation, 206  
 Accretional heating, 217  
 Accretion of the planets, 215–16  
 Acid rain, caused by impact, 223  
 Acoustic fluidization, 151–154, 177, 180  
 Aerodynamic stresses on meteoroids, 206–8  
 Agglutinate, 110, 196–97  
 Ahrens, T. J., 11  
 Airfall deposits, 18  
 ALHA 81005, 107  
 Alpine Valley, lunar, 95  
 Althans, 4  
 Alvarez, L. W., 222  
 Alvarez, W., 222  
 Ames Vertical Gun Range, 11, 49, 101, 120–21  
 Andrews, R. J., 110  
 ANEOS code, *See* Equation of state, semianalytic  
 Angle of impact, distribution, 49  
 Antipodal focusing of seismic energy, 219–20  
 Antoniadi crater, 132  
 Apollo program, 11  
 Arandas crater, 98–99  
 Archean impacts on earth, 221–22  
 Argyre basin, 172  
 Armor-penetrating projectiles, 10, 51  
 Asthenosphere, 176–77  
 Atmospheres  
 impact origin, 222  
 resistance to meteorite penetration, *See* Meteorites, penetration of atmosphere  
 Atmospheric scale height, 207–8  
 Baldwin, R. B., 6, 110, 114, 131, 141, 148, 169  
 Ballistic formulae, 87  
 Ballistic pendulum model, 11, 64  
 Ballistic sedimentation, 92–93, 100  
 Barringer, D. M., 6, 7  
 Base surge, 92–93, 100  
 Beals, C.S., 8  
 Beer, N. and J. H. Madler, 4  
 Beethoven crater, 170  
 Bernoulli Equation or Law, 53, 76  
 Bessel crater, 18, 19  
 Bingham fluid, 148–51, 153  
 approximated by a power-law fluid, 153–54, 158  
 impacts into, 149, 151  
 Birkoff, G., 10  
 Bjork, R. L., 10, 11  
 Block size on crater rims, 88  
 Blowout, atmospheric, 212  
 Boon, J. D. and Albritton, C. C., 8
- Bosumtwi crater (also known as Ashanti crater), 7, 212  
 Bow shock, 206, 208  
 Branca, W. and Fraas, E., 7, 8  
 Breakup of meteoroids in the atmosphere. *See* meteoroids, atmospheric breakup  
 Breccia lens, 17, 126–30, 141–42  
 Brent crater, 8, 122, 127  
 Bryan, J. B., 114  
 Bucher, W. H., 7  
 Butterfly wing ejecta pattern. *See* Ejecta distribution, Butterfly wing  
 Caloris basin, 95, 170, 181, 219  
 Cameron, A. G. W., 224  
 Cauchy Rille, 161  
 Cayley formation, 95  
 Central peak complex, diameter, 132, 138, Table 8.1  
 Central peaks: 18, 82, 129, 132–38, 147–49  
 formation, 136, 147–51  
 frozen Tsunami theory, 148–49  
 height, 131, 132, Table 8.1  
 hydrodynamic theory, 147–50  
 theories of origin, 141  
 Ceres, 218  
 Chapman, C. R., 195–96  
 Charters, A. C., 11, 64, 117  
 Chevron dunes. *See* Herringbone pattern  
 Chladni, E. F., 4, 6  
 Clustered impacts, 105, 209–11  
 Circularity of craters, 5, 6  
 Clasts in ejecta, 90–92  
 Coesite, 8, 41  
 Cohesion, 145, 147, 149–50, 151, 153, Table 8.2  
 Complex craters, 130–54  
 depth, 145, 146. *See also* complex craters, morphometry  
 floor, 142, 160  
 morphology, 18–22, 131–33  
 morphometry, 131–133  
 stratigraphic uplift, 136  
 Compton crater, 132, 134, 165  
 Concentric craters, 82–84  
 Contact and compression stage, 46–59  
 duration, 46, 48, 50, 58  
 energy partition, 57  
 maximum shock pressures, 54, 56  
 oblique impacts, 49–51  
 shock pressures during, 53–58  
 unloading of the projectile, 57–58  
 Continuous ejecta. *See* ejecta blanket
- Cooling  
 of breccia lens, 129–30  
 of multiring basins. *See* Multiring basins, thermal evolution  
 Copernicus crater, 18, 101, 104–5, 110, 131–33, 142, 202, 219  
 olivine in central peaks, 133  
 Cordillera Mountains of the Moon, 22, 26, 163, 173, 175, Table 9.3  
 Cosmic dust, 206  
 Coulomb friction, 34–35, 145  
 Crater collapse: 84, 88, 141–54. *See also* Modification stage  
 bingham fluid theory, 149–50  
 phenomenology, 144–50  
 physical theory, 150–54  
 stability limits, 145  
 timescale, 126, 142  
 Crater depth, apparent, 77, 112–13  
 Crater diameter, apparent, 77, 112–13  
 Crater field, 207–10  
 Crater floor, 141–42, 160, Table 8.1  
 Crater formation time, 77, 123–24  
 Crater growth, 47, 48, 76–78  
 Crater morphology  
 aberrant, 27  
 benches, 82–84  
 clustered impacts. *See* Clustered impacts  
 complex. *See* Complex craters  
 inverted sombrero, 27, 209  
 microcraters. *See* Microcraters  
 multiringed. *See* Multiringed basins  
 peak-ring. *See* Peak-ring craters  
 simple. *See* Simple craters  
 Crater obliteration, by other craters.  
*See* Crater populations, evolution  
 Crater populations, 184–96. *See also* size-frequency distribution  
 cumulative saturated density, 192  
 equilibrium, 191  
 evolution, 190–96  
 geometric saturation, 191–92  
 growth of transition diameter with time, 194  
 lunar highlands, 196  
 lunar post-mare, 187  
 saturation, 191  
 Crater profile, analytic representation, 158  
 Crater rays. *See* Rays, crater  
 Crater rim, 87–89  
 blocks on rim, 88–89  
 .. collapse, 84, 88, 126, 128–29  
 height, 88, 129, 144, 202  
 overturned stratigraphy, 9, 18, 79, 87–89

## INDEX

**Crater shape**  
 effect of an atmosphere, 209–11  
 dependence on original topography, 21, 26, 84  
 dependence on planes of weakness, 84

**Crater size.** *See* Scaling relations in oblique impacts, 50–51. *See also* Scaling relations, oblique impact

**Cratering**  
 experimental study, 11  
 numerical simulation, 10, 11

**Cratering efficiency**, 119, 121

**Cratering rate**  
 lunar, 190  
 terrestrial, 190

**Craters, circularity.** *See* Circularity of craters

**Craters, elliptical.** *See* Elliptical craters

**Craters in porous targets**, 14

**Creep of solids**, 154, 156

**Cretaceous-Tertiary impact event**, 8, 41, 70, 222–23

**Croft, S. K.**, 175

**Cryptovolcanic structures**, 7, 8

**Dana, J. D.**, 4

**Dating, limitations from crater counting**, 191, 196

**Decarbonation by impacts.** *See* Impact decarbonation

**Decaturville crater**, 7, 209

**Dence, M. R.**, 9, 127, 130

**Depth/Diameter ratio**, 78, 81, 126, 129–30, 132–33, 138, 140

**Depth of burial**  
 equivalent, 11, 113–14, 141  
 optimal, 113  
 scaled, 113–14, 120

**Depth of excavation.** *See* Excavation depth

**Detached shock,** *See* Shock wave, detached

**Dienes, J. D. and Walsh, J. M.**, 11, 116

**Dietz, R. S.**, 6, 8

**Displacements, caused by shock wave**, 62–63

**Disruption energy**, 219

**Disruption of planets by impact**, 216, 218–19

**Domed floors**, in relaxed craters, 160

**Double pulse wave structure**, 37, 40

**Droplets, size condensed from vapor**, 70–71

**Ejecta**  
 atmospheric effect on, 100, 212  
 block sizes, 104–7  
 fast, 72–74. *See also* Spalled ejecta  
 lightly shocked, 62–64. *See also* Spalled ejecta  
 lobate patterns, 98  
 provenance, 78  
 range, 87  
 shock level, 75, 76  
 time of flight, 87  
 total mass or volume, 80, 125  
 velocity, 74–76, 80–81, 92–93, 124–25

**Ejecta blanket**  
 clast distribution, 90–92  
 extent, 18, 22, 90  
 facies, 94–95  
 fluidized, 27, 95–100  
 morphology, 18, 22  
 self-similarity, 90, 124–25  
 thickness, 90, 125, 202

**Ejecta curtain**, 74–75, 92–93

**Ejecta distribution**  
 butterfly wing, 49, 101–3  
 forbidden zone, 82, 101–3  
 oblique impacts, 49, 82, 101

**Ejection angle**, 80, 92

**Elastic blocking temperature**, 182

**Elastic precursor wave**, 36, 37

**Elastic rebound**, 62

**Elastic waves**, 30–33, 62  
 energy density, 31  
 longitudinal, 30  
 particle velocity, 30  
 radiated by impact, 67–68  
 reflection at interfaces, 31–33  
 speed, 30  
 stress in, 30–31  
 tensile phase, 62  
 transverse, 30

**Elliptical craters**, 5, 6, 49–50, 81–82

**Energy coupling, projectile into target**, 48, 49

**Energy of crater formation**, 114, 121. *See also* Scaling relations, energy scaling

**Energy partition**, 57, 66–67

**Equation of state**, 11, 38–40, 57, 65, 68, 114, 230–34

**hydrocode computations**, 230

**linear shock-particle velocity.** *See* Shock velocity-particle velocity relation, linear

**liquid**, 230

**Mie-Gruneisen**, 321

**Murnaghan**, 57, 321, Table AII.1

**P-V plot** 38–39

**perfect gas**, 68, 230–31

**U-u<sub>p</sub> plot**, 38–39

**porous solids**, 40–41

**semianalytic**, 234

**stiffened gas**, 231

**Tillotson**, 39, 233, Table AII.3

**Thomas-Fermi**, 231–32

**Equations of motion, hydrodynamic.** *See* Hydrodynamic equations

**Equivalent center, of shock wave**, 71, 81

**Erosion of plants by impacts**, 218

**Escape velocity**, 215

**Euler crater**, 18, 20

**Excavation, by displacement**, 74, 78, 80

**Excavation depth**, 78, 80

**Excavation flow**, 60, 74–81  
 geometry, 74–76, 80  
 hydrodynamic regime, 62  
 in layered targets, 82–84, 141  
 initiation by shock, 74  
 mixing of material, 75, 78  
 velocity, 74, 80

**Z-model.** *See* Z-model

**Excavation stage**, 46, 47, 60–86

**energy partition**, 66–67

**Explosion craters**, 5, 6, 9, 112–13, 120  
 diameter vs. depth of burial, 113, 120

**Fillets around lunar rocks**, 184

**Fireball, radius**, 212

**Floor-fractured craters**, 161, 175

**Flow parameter**, 149

**Fluidization of rock debris**, 151–54

**Fluidized ejecta.** *See* Ejecta blankets, fluidized

**Flynn creek crater**, 7, 209

**Forsterite, liquid-vapor phase curve**, 70

**Free surface, effect on shock wave.** *See* Interference zone, near surface

**Froude number**, 116, 117

**Galileo, G.**, 3, 4

**Gassendi crater**, 161

**Gault, D. E.**, 11, 51, 64–66, 100–1, 120–22, 191–92, 194, 200–2

**Gault-Heitowit model**, 64–66

**Gifford, A. C.**, 6, 113

**Gilbert, G. K.**, 4, 5, 6, 14, 95, 163

**Gilgamesh basin**, 173

**Gilvarry, J. J. and Hill, J. E.**, 11

**Glen's law of ice flow**, 153

**Grabens**  
 concentric, 181  
 radial, 161

**Grady-Kipp fragments**, 64, 71, 107

**Gravitational binding energy**, 219

**Gravity, effect on crater size.** *See* Scaling relations, gravity dependence

**Gravity anomalies over craters**, 18, 22, 127, 140

**Greeley, R.**, 100

**Grieve, R. A. F.**, 9, 127

**Grimaldi crater**, 176, 181

**Gruithuisen, 4**

**Gruneisen parameter**, 39, 231

**Hager, Dorsey**, 7

**Halo craters**, 101

**Hankel transform**, 158

**Harriot, T.**, 4

**Hartmann, W. K.**, 163, 169, 197, 224

**Head, J. W.**, 170, 175

**Heat, deposited by impacts**, 217

**Heating by shock waves.** *See* Shock heating

**Heitowit, E. D.**, 64–66

**Hellas basin**, 172

**Henbury craters**, 7, 208–9

**Herringbone pattern**, 95, 101, 104

**Hevelius formation**, 26

**Holsapple, K. A.**, 11, 119, 121

**Hooke, R.**, 4

**Huascaran avalanche**, 93

**Hugoniot**  
 curve, 38–43  
 equations, 37  
 equations, derivation, 228–29  
 elastic limit, 34–35, 39, 64, 72, Table 3.1

**Hugoniot, P. H.**, 37

**Humboldt crater**, 161

**Hummocky terrain**, 94

- Hydrodynamic equations, 114  
 Hydrodynamic flow, of projectile, 47–48  
 Hydrodynamic similarity, 114–15
- Ice calderas, 175  
 Imbrium basin, 78–79, 95, 221  
 Imbrium sculpture, 4, 95  
 Impact decarbonation, 222, Table 3.3
- Impact**  
 low velocity, 5, 14  
 –explosion analogy, 5–7, 11, 48, 113–14  
 depth of burial. *See Depth of burial, equivalent*  
 melt, 18, 75–76, 78, 91–92, 122–23, 127–29, 142, 173. *See also Suevite*  
 melt in porous rocks, due to jetting, 53  
 melt pool, 127–28  
 melting, 44, 63–64, Table 3.3  
 vaporization, 44, 63–64, 112–23, Table 3.3
- Impedance matching.** *See Planar impact approximation*
- Interference zone**  
 near-surface, 71–74, 107  
 depth, 72  
 stress wave, 71, 72
- Iridium, in Cretaceous-Tertiary clays, 222
- Isidis basin, 172
- Isostatic equilibrium, 160
- Ives, H. E., 5, 9
- Ivory coast tektites, 212
- Jangle U nuclear explosion crater, 9, 87
- Jefferson, Thomas, 6
- Jet penetration depth, 114
- Jetting, 10, 46, 48–53, 224–25  
 amount of material in jet, 51, 53, 224  
 asymmetric, 52  
 coordinate frames, 52  
 critical angle, 51  
 discovery, 51  
 duration, 53  
 jet velocity, 52  
 jet width, 53  
 pressure in jet, 53  
 symmetric, 51–53
- Joints, effect on crater shape, 84
- Kapoeta meteorite, 184
- Kentland structure, 7, 8
- King crater, 26, 27, 84
- KOA nuclear crater, 151
- Kuiper G. P., 6, 163
- Lampson, C. W., 9, 112
- Lampson scaling, 9, 112–13, 115
- Late heavy bombardment, 190, 196, 221
- Late stage equivalence, 11, 116–17
- Layered targets, effect on craters, 82–84
- Linné crater, 14, 16
- Lithosphere  
 definition, 176
- effect on viscous relaxation, 161  
 elastic, 161  
 thickness, 179–81  
 yield strength, 179
- Lonar crater, 127
- Lowe, D. R., 221
- Low-velocity impacts, 5, 215–16
- Lunar craters  
 circularity, 5, 6  
 impact origin, 4, 5, 6  
 volcanic origin, 4
- Lunar soil. *See Regolith*
- Mach number, 206
- Mckinnon, W. B., 66, 176
- Manicouagan crater, 8, 170, 215
- Mantle plug, beneath basins, 160, 167–68
- Mare ridges, 181
- Mascon**  
 definition, 180  
 tectonics, 180–81
- Maskelynite, 41
- Maxwell, D. E., 10, 80–81
- Maxwell time, 156, 176
- Megablocks, of Ries crater, 105, 107
- Meen, V. B., 8
- Megaregolith, 197
- Melt or Melting.** *See Impact Melt or Melting*
- Melting, pressure-relief, 173, 221
- Melt pool, in bottom of craters, 127–28
- Meteor crater, AZ, 6, 7, 8, 16–18, 41, 70, 84, 87, 91, 114, 126–27, 136, 219
- central mound, 136  
 formation energy, 114  
 mass of projectile, 7, 16  
 tear faults, 84
- Meteor flight through the atmosphere, 205–7
- Meteorite hazard, 9, 11
- Meteorites  
 chondrites in lunar soil, 73  
 cosmic ray exposure ages, 218  
 discovery, 6  
 entry velocity, 205  
 Euclites, 216  
 lunar, 73, 107  
 minimum size to penetrate atmosphere, 205, Table 11.1  
 penetration of atmosphere, 6, 7, 205–11  
 SNC, 68, 107
- Meteorid flux, 9, 187–90
- Meteoroids**  
 atmospheric breakup, 207–9  
 heating in atmosphere, 206
- Meteors  
 energy deposition in atmosphere, 207  
 visual, 207
- Microcraters, 14, 15, 116
- Micrometeoroids, 206
- Microtektites, 212, 222
- Mie-Gruneisen equation of state.** *See Equation of state, Mie-Gruneisen*
- Mimas, craters on, 190, 192
- Modification stage, 47, 126–62
- Moldavite tektites, 212
- Molecular flow regime, 205–6
- Momentum in shock wave, 62
- Moon, origin by impact, 223–26
- Moulton, F. R., 7
- Mudflows, viscosity, 151
- Multiring basins, 163–83  
 asymmetric rings, 165, 167  
*on Callisto, 168–69, Table 9.1. See also Valhalla basin*  
 discovery, 163  
 furrows on Ganymede, 168  
*on Ganymede, Table 9.2, 169*  
 megaterrace theory, 175  
*on the moon, Table 9.3. See also Orientale basin, Imbrium basin morphology, 22–24, 26–27  
 nested crater theory, 82, 84, 175–76  
 occurrence, 169–173  
 ring spacing, 169  
 ring tectonic theory, 176–80  
 sinuous ridges, 168  
 symmetric rings, 165  
 thermal evolution, 182–83  
 Valhalla type, 168–69  
 volcanic modification theory, 173–74*
- Murnaghan equation of state.** *See Equation of state, Murnaghan*
- Mushroom cloud, 212  
 formation timescale, 212
- Nasmyth and Carpenter, 4
- New Quebec crater, 8
- Newporte crater, 215
- Near-surface zone.** *See Interference zone, near-surface*
- Nectaris basin, 95
- Nininger, H. H., 70
- NO**  
 creation by shock, 222  
 role in Cretaceous extinction, 223
- Nordyke, M. D., 10, 120
- Nuclear explosions, maximum yield, 212
- Nuclear winter, 223
- Oberbeck, V. R., 82, 93, 101, 199, 200
- Oblique impact,** 49–51, 81–82, 101, 121–22
- Oblique impact craters**  
 morphology, 21  
 size. *See Scaling relations, oblique impact*
- Odessa, Texas crater, 7, 87
- Oil production from craters, 215
- O'Keefe, J. D., 11
- Operation Peggy, 180
- Öpik, E. J., 5, 10, 63
- Optimum depth of burial. *See Depth of burial, optimal*
- Orbital injection of impact ejecta, 224
- Orientale basin, 22, 23, 26, 95, 163–68, 175, 221
- ejecta, 22, 26, 95, 101, 173
- Origin of the moon, 224. *See also Moon, origin by impact*
- Overturned stratigraphy. *See Crater rim, overturned stratigraphy*

## INDEX

- Palimpsest crater, 156, 160  
 Pan crater, 126  
 Particle velocity  
     attenuation in shock wave, 62–63, 64–68  
     definition, 29  
     in wave. *See* wave type: pressure, elastic, plastic, shock, etc.  
 Peak-ring craters, 21, 132, 138, 143, 149, 151, 163, Table 8.1  
 Pedestal craters, 98  
 Penepalimpsest crater, 156  
 Penetration depth. *See* Jet penetration depth  
 Petrographic shock indicators, 41, 42, Table 3.2  
 Phase transformations in shocks, 40–44  
 Pieters, C., 133  
 Pit craters, 21, 133, 139  
 Pi-theorem, 117  
 Planar Impact Approximation, 11, 54–57  
     algebraic equation, 56  
     graphical method, 55  
     limitations, 58  
 Planetary collisions, 216, 222–26  
 Planetesimals, 215–16, 218–19  
 Planets, thermal regime during growth, 216–18  
 Plastic failure law, 34, 145, 147  
 Plastic waves, 33–37  
     nonlinearity, 37  
     particle velocity, 36, 39  
     speed, 36, 39  
     stress, 35  
 Poisson ratio, 31  
 Porosity  
     effect on crater size. *See* Scaling relations, porosity dependence  
     effect on Hugoniot curve, 40, 41  
 Prairie flat crater, 151  
 Pressure attenuation in shock wave, 62–68  
 Pressure waves, 29, 30  
     energy density, 29  
     particle velocity, 29  
     speed, 29  
 Proctor, R. A., 4  
 Production population, definition, 190  
 Quaide, W. L., 82  
 Quartz, 40  
     shocked, 8, 41, 222, Table 3.2  
 Radial fractures around craters, 161  
 Rampart craters, 95, 100  
 Rarefaction wave. *See* release wave  
 Rayleigh line, 39, 41, 43, 44  
 Rays: crater, 4, 107–10  
     extent, 110  
     nature, 110  
     obliteration, 108, 110  
 Red Wing Creek crater, 215  
 Regolith, 184, 196–203  
     contamination by meteorites, 196  
     cosmic ray exposure time, 201  
     gardening, 108, 184, 200–1  
     growth, 197–200  
     horizontal transport, 201–3  
     maturity, 196–97  
 Monte Carlo model, 199–200  
     nonlinear growth rate, 198–200  
     overturn time 200–1  
     Shoemaker model, 197–99  
     thickness estimates 82, 197–200  
 Relaxation time, viscous, 156, 158–59  
 Release adiabats, 39, 43, 44  
 Release wave, 42, 48, 57  
     particle velocity, 42  
     speed, 42, 57  
     tensile phase, 71, 72  
 Residual velocity, 42, 43  
 Reynolds number, 116  
 Rheologic models of planetary surfaces, 158–61, 177  
 Ricochet of projectile, 49  
 Riemann invariant, 42  
 Ries crater, 8, 41, 88, 91–93, 105, 107, 140–41, 151, 212  
 Rim. *See* Crater rim  
 Ringwood, A. E., 224  
 Rook Mountains of the moon, 22, 26, 163–67, 175  
 R-plot. *See* size-frequency distribution, R-plot  
 Runaway growth of planets, 216  
 Safronov number, 215  
 Saturation. *See* crater populations, saturation  
 Scaling relations, 9, 10, 112–25  
     Charters and Summers' relation, 117  
     coupling constant approach, 117, 119, 123–24  
     crater diameter, 119, 120–21  
     crater volume, 119  
     ejecta, 124–25  
     energy scaling, 113–15, 120–21  
     Gault's relations, 120–21  
     gravity dependence, 115, 118–19  
     melt and vapor volumes, 122–23  
     oblique impact, 120–22  
     Pi-group approach, 117–21  
     porosity dependence, 118  
     power-law, 119  
     Schmidt-Holsapple, 118–19, 121  
     similarity, 112–15, 117–18  
     strength dependence, 116–17, 119–20  
     strength-gravity transition, 116, 119–20  
     viscosity dependence, 116  
     yield scaling, 112–13, 120–21  
 Scatter ellipse, 208–9. *See also* Crater field  
 Schiller crater, 25  
 Schmidt, R. M., 10, 11, 119, 121  
 Schrödinger crater, 95, 165  
 Schröter, J. H., 4, 90  
 Schröter's rule, 90  
 Schultz, P. H., 100  
 Secondary craters, 94, 95, 101–7  
     morphology, 14, 16, 104–5  
     size, 101  
 Secondary crater fields, 101, 104–7  
     density, 105–6  
     extent, 22, 100, 105–6  
     mass of impactors, 106–7  
 Seismic efficiency, 67  
 Seismic magnitude of impact, 67, 219  
 Seismic modification of topography, 219–21  
 Seismic waves, generated by impact, 66–67, 219, 221  
 Serpent mound structure, 7  
 Shatter cones, 8, 42, 138, 140  
 Sherman glacier landslide, 95  
 Shock front, thickness, 37, 61, 66  
 Shock heating, 44, 65  
 Shock metamorphism, 40–44  
 Shock velocity-particle velocity relation, linear, 39, 232–33, Table AII.2  
 Shock wave, 37–42  
     detached, 11, 58, 60, 61  
     energy, 62, 65  
     energy density, 37  
     experimental study, 38  
     geometry in impacts, 48, 50, 60–62  
     irregularity, 41, 42, 74, 154  
     particle velocity, 37, 39  
     rise time, 61, 71  
     scattering by inhomogeneities, 154  
     speed, 37, 39  
     stress in, 37, 38  
     thermodynamic irreversibility, 39, 41, 43  
 Shock wave expansion and attenuation, 60–68  
 Shoemaker, E. M., 8–11, 49, 87, 88, 114, 196–200  
 Sierra Madera crater, 7, 136, 140  
 Simple craters, 14–18, 126–30  
     gravity anomaly over, 127  
     formation, 126–30  
     morphology, 14, 16–18, 126  
 Simple-to-complex transition, 21, 130–31  
 Sinus Medii, crater population, 194  
 Size-frequency distribution, 186–87  
     cumulative, 186–87  
     fractional area covered, 187  
     incremental, 184  
     power-law, 187  
     R-plot, 187  
 Size-morphology progression, 4, 5, 15, 27, 170  
 Slump terraces. *See* Terraces, slump  
 Slumping, 130. *See also* Complex craters  
 Smooth plains, 95  
 SNC meteorites. *See* Meteorites, SNC  
 Snowball crater, 177  
 Spall, near impacts, 71–74, 78  
 Spalled ejecta. *See also* Interference zone  
     ejection time, 73  
     velocity, 72–73  
     volume, 73, 107  
 Spall plates, 14, 72, 107  
 Speed of waves. *See* under wave type; e.g., pressure, elastic, plastic, shock, etc.  
 Spherule beds, Archean, 221  
 Spherules. *See also* Droplets  
     in lunar soil, 71, 222  
     nickel-iron, 70  
     Sanidine, 222  
 Stages of impact cratering, 46, 47  
 Steinheim basin, 7, 8  
 Stishovite, 8, 40, 41

- Straight wall, 161  
Strength, effect on crater size. *See* scaling relations, strength dependence  
Strength crater, 175  
Strength parameter, 149  
Stress, beneath a crater, 144, 145, 179  
Stress wave, *See* shock wave  
Strewn field. *See* Crater field  
Structural control of crater shape, 84  
Stagnation pressure, 206  
Sturzstroms, 150  
Sudbury crater, 170, 215, 221  
Suevite, 92, 140–42, 173. *See also* Impact melt  
Tait, P. G., 147  
Teapot Ess nuclear crater, 9, 10  
Tektites, 70, 212  
Tensile fracture, 71–72  
Tensile strength, dynamic, 72, 107, 218  
Terminal velocity, 206  
Terraces, slump, 4, 5, 18, 84, 142–43, 175  
width, 142–43, 147  
Terracettes, 184  
Terrestrial craters, recognition, 6–9  
Theophilus crater, 18, 21  
Thermal effects, of simple craters, 129–30  
Thomas–Fermi limit. *See* Equation of state, Thomas–Fermi  
Tilghman, B. C., 7, 9  
Tillotson equation of state. *See* Equation of state, Tillotson  
Timocharis crater, 89  
Tolstoj crater, 170  
Topography, seismic modification of. *See* Seismic modification of topography  
Toreva blocks, 142–43  
Transient crater, 76–77, 112  
depth/diameter ratio, 78, 81, 129, 138, 140, 143–44  
formation time, 77, 124  
shape, 76–78, 143–44  
Tsilokovsky crater, 95  
Tunguska explosion, 7, 207, 222  
Tycho crater 107–8, 202  
Urey, H. C., 6  
Valhalla basin, 22, 24, 26, 27, 168, 179  
Van Dorn, W. G., 148  
Vapor plume, interaction with an atmosphere, 211–13  
Vaporization. *See also* Impact vaporization  
heat of vaporization, 69, 70, 122  
of projectile, 5, 7, 48, 51  
Vapor plume, 58, 68–71  
condensation, 70–71  
entrainment of ejecta, 70  
velocity of expansion, 58, 69  
Zel'dovich and Raizer model, 69  
Velocity doubling rule, 32, 43  
Viscosity, 116, 147, 149, 151, 153, 156, 176, 180, Table 8.2  
Viscous degradation of craters, 154–61  
effect of lithosphere, 161  
Volcanism, impact-induced, 221  
Vredefort structure, 7, 170, 221  
Waste heat, in shock compression, 65  
Water table  
depth on Mars, 100  
effect on craters, 82, 84, 100  
Wave speed. *See under* wave type; e.g., pressure, elastic, plastic, shock, etc.  
Wedekind, J. A., 101  
Wegner, A., 5  
“Weird terrain,” 219  
Wetherill, G. W., 215, 217  
West Hawk Lake crater, 127  
Yuty crater, 98  
Z-model, 10, 77, 80–81  
numerical tests, 81  
Zond 8, 163

