INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0800

UMI®
As members of the Final Examination Committee, we certify that we have read the dissertation prepared by JAMES NORMAN HEAD entitled FRAGMENTATION AND EJECTION OF THE MARTIAN CLAN METEORITES and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

[Signatures and dates]

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College. I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

[Signature and date]
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: [Signature]
ACKNOWLEDGEMENTS

I thank Jay Melosh, who has provided support to me as well as to this work whenever I was wise enough to ask for it, even in those days when I was not a good risk. He has been most helpful in assessing my career opportunities as well. Jay has been lauded often as a great if demanding advisor, an outstanding educator, and top notch scientist in just about anything that catches his attention. Certainly I can add to that litany. However, in the adjectival pantheon surrounding this man—wander the halls of Gilruth some fine Sunday night in March for a sampling—my contribution would have to be "compassionate."

I thank Boris Ivanov for his willingness to lend his experience and brilliance in making the completion of this project possible. It was his insight that pointed the way toward a simple method of simulating impacts in high resolution with SALE. Without that capability, this problem would remain unsolved (see Figure 5.1). I have been most fortunate to have the assistance of these two scientists, both Barringer Medal winners, during this work.

I acknowledge the brief collaborations I had with Bob Singer and Tim Swindle before I considered computer modeling for my thesis work. The depth in which I learned about remote sensing and geochemistry made it possible to put my research into the broader context of scientific knowledge, and hence made it much easier to recognize the worth of my results. Had I plunged headlong into impact physics, surveying the martian cratering and meteorite geochemistry literature to the point that I could understand what I accomplished would have been truly daunting.

I acknowledge Carl Sagan and the producers of Cosmos for introducing me to planetary science at an impressionable age, and to Gene Shoemaker for his role in founding the field of impact cratering.

Finally, a nod to Stella at Infotrieve for hunting down the most obscure references, including one that had eluded the impact physics group at LPL for 2.9 dissertations.
DEDICATION

A time of joy, a time of sorrows. A time of growth, a time of anguish. A time of trouble, a time of triumph. All this and more a graduate school education so often is. Coming of age events are so often the most awkward of situations, and so many of these moments have transpired in the decade (!) now ending. The best of these was 15 February 1998 (the wedding); the worst I will not recall in detail.

I must thank my parents, especially my dad, for supporting my life in science. This started about the time that I learned to read and encompassed such highlights as telescopes, microscopes, a lunar globe (now out of date) and paying my first college tuition bills at age 12. That banner year introduced me to Richard Francis and Ross Iverson, my 7th grade music and science teachers respectively. To them I owe both my profession and my principle avocation.

I thank my college mentor, Bill Graham, who has guided so many young scientists into professional life. He entrusted his irreplaceable Varian EPR Spectrometer to an undergrad who didn't even like quantum mechanics that much and provided me with my first opportunity to teach undergraduate astronomy. He showed me then and now that science is a wonderful life path.

Jay talks of intellectual horsepower, and while a necessary condition for a successful graduate career, it is hardly sufficient. While the gentle days come not, the acceptance of comrades, even of those not in arms, upholsters one's life and nurses the soul. Of these I single out the groomsmen–Dave Mason, of surprising eloquence, Mark Lemmon, of political insight, Maryann DeMatthews, learned both in outrage and acceptance, and Richie Saunders, fellow purveyor of puns. There are many others of course, and you know who you are. Dave deserves special recognition for having to follow me through school. A nod to Randy Oberhoff and Fred Wiggins for keeping the marbles secure.

I want to mention my three relatives who always supported my passion for learning but who did not survive my tenure in graduate school. First my grandmother Tina Isobel Head, who sat up with me all night to count the stars. Second, my Aunt Helen, who walked with me to the light-free if not-that-safe parts of St. Louis to watch the moon and Venus set together (I was still too young to know better). Finally her husband my uncle Gene Walker, who spent his 18th year in the European Theater of Operations, ensuring that my generation could spend their 18th year (and beyond) in less arduous pursuits. He always wanted to know what we learned from all those moon rocks, and I am thankful he lived long enough to hear the tale.

And always, for Molly, the sine qua non.
TABLE OF CONTENTS

LIST OF FIGURES .......................................................... 8
LIST OF TABLES ............................................................. 10
ABSTRACT ........................................................................ 11
1 Introduction ................................................................... 12
2 The SNC Provenance Problem ........................................... 16
   2.1 The Mars of the Imagination ........................................... 17
   2.2 Rocks from Mars? .......................................................... 18
      2.2.1 Why the SNCs are martian ........................................ 20
      2.2.2 The physicists respond ............................................ 23
      2.2.3 Loophole–interference with the free surface ............... 24
      2.2.4 Stress waves in a 1–dimensional rod ......................... 27
   2.3 Hydrocode Modeling of Impacts .................................... 31
   2.4 Interdisciplinary Conflict .............................................. 34
   2.5 Conclusion ................................................................. 36
3 The Martian Clan Meteorites ............................................. 48
   3.1 Review of Martian Meteorite Geochemistry ................. 48
      3.1.1 Shergottites ........................................................... 50
      3.1.2 Nakhlites .............................................................. 51
      3.1.3 Chassigny .............................................................. 51
      3.1.4 ALH84001 ............................................................... 52
   3.2 Martian Meteorite Ages ............................................... 52
      3.2.1 Crystallization ages ............................................... 53
      3.2.2 Crater ages ........................................................... 54
      3.2.3 Cosmic ray exposure ages ....................................... 57
   3.3 Delivery of Rocks from Mars to Earth ......................... 59
   3.4 Delivery Timescales .................................................... 61
      3.4.1 Opik methods ........................................................ 62
      3.4.2 Direct numerical integration .................................... 63
   3.5 Delivery of Martian Material to the Earth ...................... 64
      3.5.1 Terrestrial ages and in-space breakup ...................... 65
   3.6 Conclusion ................................................................. 68
4 The SALE 2D Hydrocode ................................................. 71
   4.0.1 Why hydrocodes? .................................................... 72
   4.1 SALE ......................................................................... 73
      4.1.1 Dynamic fracture ................................................... 74
      4.1.2 Weibull distribution ............................................... 75
      4.1.3 Grady-Kipp fragmentation model ............................ 77
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS — Continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Upgrades to SALE</td>
</tr>
<tr>
<td>4.2.1 Multiple material code</td>
</tr>
<tr>
<td>4.2.2 Alternate node coupler</td>
</tr>
<tr>
<td>4.2.3 Strain calculation</td>
</tr>
<tr>
<td>4.3 High Resolution in the Spall Zone</td>
</tr>
<tr>
<td>4.3.1 Two step solution</td>
</tr>
<tr>
<td>4.4 Conclusions</td>
</tr>
<tr>
<td>5 Spallation</td>
</tr>
<tr>
<td>5.1 Analytic Spall Models</td>
</tr>
<tr>
<td>5.2 Measuring the Interference Zone</td>
</tr>
<tr>
<td>5.3 Parameter Studies</td>
</tr>
<tr>
<td>5.3.1 Resolution studies</td>
</tr>
<tr>
<td>5.3.2 Artificial viscosity</td>
</tr>
<tr>
<td>5.3.3 Weibull parameters</td>
</tr>
<tr>
<td>5.3.4 Material properties</td>
</tr>
<tr>
<td>5.4 Layering Effects</td>
</tr>
<tr>
<td>5.4.1 Alteration of the shock wave</td>
</tr>
<tr>
<td>5.4.2 Spall velocity vs. distance</td>
</tr>
<tr>
<td>5.5 Spall Thresholds</td>
</tr>
<tr>
<td>5.5.1 Competent material</td>
</tr>
<tr>
<td>5.5.2 Damaged material</td>
</tr>
<tr>
<td>5.5.3 Speculations on porosity</td>
</tr>
<tr>
<td>5.6 Conclusion</td>
</tr>
<tr>
<td>6 Scenarios for Martian Meteorite Launch and Delivery</td>
</tr>
<tr>
<td>6.1 Impacts into Fresh Material</td>
</tr>
<tr>
<td>6.2 Impacts into Weathered Material</td>
</tr>
<tr>
<td>6.3 Discussion</td>
</tr>
<tr>
<td>6.4 Conclusion</td>
</tr>
<tr>
<td>7 Summation</td>
</tr>
<tr>
<td>7.1 What has been Resolved</td>
</tr>
<tr>
<td>7.2 What Remains to be Done</td>
</tr>
<tr>
<td>REFERENCES</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

2.1 Nonphysical isobars from an impact ........................................... 26
2.2 Physically plausible isobars ...................................................... 27
2.3 Accelerometer data and interpretation from a nuclear test .............. 28
2.4 Impingement of compressive wave on the free surface .................... 29
2.5 Stress wave in a one-dimensional rod .......................................... 37
2.6 Velocity doubling in a one-dimensional rod .................................... 38
2.7 Sound speed in Westerly Granite ................................................. 39
2.8 Pressure wave in a one-dimensional rod ....................................... 40
2.9 Interference with the free surface .............................................. 41
2.10 Reflection of a tensile wave ..................................................... 42
2.11 Damage in a one-dimensional rod ................................................ 43
2.12 Maximum pressure in a one-dimensional rod .................................. 44
2.13 Damage propagation ............................................................... 45
2.14 Fragment sizes in the one-dimensional rod ..................................... 46
2.15 Laboratory demonstration of spall .............................................. 47
3.1 Lunar impactor flux curve ......................................................... 69
3.2 Ejection ages for martian meteorites .......................................... 70
4.1 Fragment size distributions: experiments vs. numerical model ........... 86
4.2 Initial Eulerian calculation ......................................................... 87
4.3 Initial grid, Lagrangian calculation ............................................. 88
4.4 Computational grid at 11.5 msec ................................................ 89
4.5 Computational grid at 23 msec ................................................... 90
4.6 Computational grid at 34.5 msec ............................................... 91
4.7 Velocity vectors at 11.5 msec .................................................... 92
4.8 Velocity vectors at 23 msec ....................................................... 93
4.9 Velocity vectors at 34.5 msec .................................................... 94
4.10 Pressure contours at 11.5 msec ................................................ 95
4.11 Pressure contours at 23 msec ................................................... 96
4.12 Pressure contours at 34.5 msec ................................................ 97
4.13 Isotherms at 11.5 msec .......................................................... 98
4.14 Isotherms at 34.5 msec .......................................................... 99
4.15 Damage at 11.5 msec .............................................................. 100
4.16 Damage at 23 msec .............................................................. 101
4.17 Damage at 34.5 msec .............................................................. 102
4.18 Contours for maximum pressure ................................................. 103
5.1 Resolution study ................................................................. 108
5.2 Peak pressure with depth for different runs ................................... 110
5.3 Peak pressure at high resolution ................................................. 112
5.4 Peak pressure with strong artificial viscosity .................................. 113
5.5 Peak pressure at lower resolution ............................................... 114
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>Shock wave in homogeneous material</td>
<td>120</td>
</tr>
<tr>
<td>5.7</td>
<td>Shock wave in layered material</td>
<td>121</td>
</tr>
<tr>
<td>5.8</td>
<td>Spall velocities in layered terrain</td>
<td>122</td>
</tr>
<tr>
<td>5.9</td>
<td>Spall velocity and layer thickness</td>
<td>124</td>
</tr>
<tr>
<td>5.10</td>
<td>Spall velocity and damage</td>
<td>125</td>
</tr>
</tbody>
</table>
# LIST OF TABLES  WDS

2.1 Petrographic shock indicators for various minerals .................. 25  
3.1 Physical data for the martian clan meteorites ....................... 49  
3.2 Selected properties of the known martian meteorites ................. 58  
3.3 Cosmic ray exposure and terrestrial ages of martian meteorites .... 60  
3.4 Delivery efficiency of martian meteoroids .......................... 64  
5.1 Weibull parameters used in this work ................................ 115  
5.2 Tillotson parameters used in this work ............................... 116  
5.3 Fragmentation data for 3.1 km crater event .......................... 128  
5.4 Fragmentation data for 3.9 km crater event .......................... 129
I have used the SALE2D hydrocode to study spall in impacts into layered terrains. Application of my results to the problem of martian meteorite provenance resolves two outstanding paradoxes. First, the minimum size crater previously thought to be required to eject martian meteorites is so large (12 km) that it is highly unlikely such an event occurred on shergottite age terrain in the last few million years. The geochemical evidence supports four launch events. This issue I have resolved by establishing a new lower limit to the minimum size crater of 3 km.

Second, the martian meteorites are dominated by shergottites (62%) which come from the youngest and apparently rarest martian terrains. The vast majority of Mars appears to be under represented. This paradox lies on the false premise that all terrains are equally efficient in launching material during an impact. I have found that the presence of a weak, low density layer suppresses spall velocity and increases shock pressures in an impact. Since the regolith on Mars can be expected to be largely impact-generated, the older terrains are covered by a greater depth of regolith. Qualitatively, older terrains are under represented in the martian meteorites because they require larger (rarer) impacts to launch material into space. I have shown this quantitatively for shergottites, nakhlites, and Chassigny. An extension of my work provides some constraints on the extent of martian ancient terrain.
CHAPTER 1

Introduction

One of the truly remarkable results of the last twenty years of planetary science is the realization that we have in our possession rock samples from the planet Mars. Amazingly, all of this material (about 80 kg) was delivered to Earth free of charge in the form of meteorites. The SNC meteorites, so-called for the type meteorite in each sub-class: Shergotty, Nahkla, and Chassigny, had long been recognized as peculiar. By 1981, a great deal of circumstantial evidence had been compiled leading to the martian origin hypothesis for these stones (Wood and Ashwal, 1981). This hypothesis was bolstered by the discovery and analysis of trapped atmospheric gases in one of the SNCs (EETA79001), the development of a theoretically plausible launch mechanism, and the identification of meteorites of unequivocally lunar origin. By the mid-1980s the martian origin of the SNC meteorites was largely settled (Wood and Ashwal, 1981; McSween Jr., 1985; McSween Jr., 1994). In retrospect, it now seems obvious that we should have expected to find material from other planets lying about on the Earth. This was even suggested in print more than 35 years ago (Shoemaker et al., 1963) and implicit in prominent models of meteorite origin and formation. The recognition of the martian origin of the SNCs has had far ranging consequences. Besides illustrating a small but very important gap in our understanding of impact physics, the notion that material can be exchanged between planets has profound implications for planetary quarantine and origin of life studies.

The qualitative version of events for the origin of the SNCs has been reasonably well-established. However, there have been inconsistent results amongst the
various scientific disciplines that have been brought to bear on the problem. In sum, it has not been possible to reconcile quantitatively the results from geochemistry with our understanding of impact physics. It is these disconnects that point the way towards achieving a better understanding of the natural world. In this particular case, one measure of our ignorance is the agreement, or rather, lack of agreement on the number of impact events required in any scenario for the formation and delivery of the SNCs to Earth. Various estimates over the years have ranged from 1 to 8 impacts. The lower estimates are driven by then state-of-the-art understanding of impact physics. It was thought that fairly large craters were required to eject the material from Mars. Since the recurrence interval for such events is large compared to the launch timescale defined by other means, many workers were driven to models using as few impacts as possible so as to be able to appeal to the statistics of small numbers. In this way they could explain otherwise irreconcilable differences between the impact physics community and the geochemists as a statistical fluke.

It is the chief result of this work that these previously irreconcilable differences have been largely resolved. Advances in computational speed and methods allow simulations of impacts and modeling of dynamic fragmentation with unprecedented spatial resolution. As a direct consequence, the results from numerical modeling of impacts show (for the first time) that the number of launching events required for the shergottites on purely geochemical grounds is the same (within errors) as the number of impacts one would expect when my results are applied to what is known about the current impactor flux at Mars. Plausible assumptions about the martian surface permits a quantifiably viable solution for the Nakhlites and Chassigny as well. Only ALHA84001 remains something of an enigma, a characterization not unique to this work.

The fact that there is now agreement for most of the problem of martian meteorite provenance indicates that this approach is likely correct and will bear further fruit as better constitutive models for more materials, especially porous materials, become available. This should help resolve the remaining discrepancies
for the SNCs. It should provide insight into issues outstanding with the lunar meteorites as well. Finally, I put to rest the notion that individual source craters for the SNCs are likely to be identified from remotely sensed data. At a conservative estimate, there are more than 13,000 craters on the surface of Mars large enough to mark the launch sites for the material we see in our collections. Moreover, the sometimes popular reliance on oblong as opposed to circular craters is shown to be of little merit since my results are for vertical impacts which of course leave circular craters. I believe that identification of specific source craters for individual SNC meteorites will require in situ analysis by teams of geologist roving the martian countryside with rock hammers, hand lenses, and mass spectrometers. The actual discovery and description of a SNC source crater will no doubt be the worthy subject of another doctoral dissertation in some not-too-distant age.

As for this dissertation, Chapter 2 reviews what is known about SNC meteorite provenance, defines the mismatch between geochemistry and impact physics, and describes hydrocode modeling of impacts. This includes a history of how we came to know that the SNCs were from Mars. Chapter 3 reviews what is known of SNC geochemistry and the most recent developments in applied celestial mechanics. These two fields provide constraints on the ejection and delivery of SNC material to earth. This in turn provides important clues as to the required size of the impact crater responsible for ejection. Celestial mechanics further constrains SNC delivery since it speaks to the relative likelihood of in-space breakup of ejected fragments. Chapter 4 reviews relevant aspects of the hydrocode I use in this work such as the calculation of fragment sizes, ejection velocities, and possible numerical artifacts. The many changes that have been made to the hydrocode since it was first used to study fragmentation are explained here. Finally Chapter 4 explains in detail how I achieved high spatial resolution in the spall zone.

The balance of the dissertation concentrates on my substantive results. Chapter 5 is dedicated to the spall phenomenon, how it is studied on theoretical grounds (both analytically and numerically), and in its applications to the present
problem. Many of the relevant parameters are identified and their influences on the results described in detail. Chapter 6 presents the most recent (and to date only quantitatively viable) scenarios for the ejection of SNC material into space and its delivery to the Earth. Chapter 7 summarizes what has been resolved and what issues remain to be settled.

The principal new result of my work is that for the first time, detailed simulations of impacts make predictions with respect to the SNCs that are quantitatively consistent with what is known from geochemical analysis and celestial mechanics. I have succeeded in simulating an impact with sufficient detail to resolve the physics dominating the spall zone from numerical artifacts. This permits the calculation of sizes, velocities, and shock state for the fragments ejected by the impact. This has allowed me to calculate the smallest crater required to eject this material to Mars' escape velocity. In addition, I have modeled the target geology, noting how layers with different acoustic properties affects the propagation of the shock wave, and hence, fragment ejection. This particular achievement allows analysis of impact conditions that cannot be modeled analytically.

The SNC provenance problem is not entirely settled, however, since there are relevant geologic materials for which good constitutive models do not exist. The chief problem here is developing a good model for porosity. It is reasonable to assume that highly weathered terrains on Mars will be relatively porous, and that this will greatly influence the spall process. Once this issue is resolved, it should be possible to model the ejection of older martian rocks (especially ALH84001) and lunar rocks as well. In summary, my work resolves a significant part of a fifteen-year-old puzzle, and in so doing indicates a likely means of solving the remaining parts as well.
CHAPTER 2

The SNC Provenance Problem

A thorough understanding of SNC provenance lies at the nexus of many diverse scientific disciplines. Among these are petrology, isotopic geochemistry, impact physics, celestial mechanics, noble gas geochemistry, photogeology, cratering statistics, and numerical modeling. In the twenty years since the scientific community first suspected that the SNCs were martian, these disparate fields of study have been brought to bear, sometimes with contradictory results. The qualitative picture has been clear for some time. The SNCs are rocks that formed on Mars hundreds of millions to billions of years ago. Somehow they were blasted from the surface and roamed through the inner solar system as meter-sized or smaller rocks for hundreds of thousands to millions of years. At various times in the last 200 ka, they fell to earth. The petrology, isotopic, and noble gas geochemistry of these rocks tie them to a large planet, which is almost certainly Mars. Mars has abundant craters, many of which are viable candidates for launch sites for these meteorites. Theoretical and experimental studies of impact physics show that intact material can be ejected to high velocities, even in excess of Mars' escape velocity. Celestial mechanics shows how such ejected material can quickly be placed into earth-crossing orbits. The difficulty is that until now, the quantitative understanding of this problem has defied solution. The results from the disparate fields listed above do not all agree. My contribution to this field is that a more detailed analysis of impact physics updates the results from that discipline in such a way as to resolve many of the discrepancies. I will devote much of this chapter to a brief review of the SNC provenance problem, thereby defining the gap in our understanding which my research fills.
2.1 The Mars of the Imagination

Mars has been of great interest to astronomers since Schiaparelli's observations in the 1880s. In particular, it was the interpretation of the observed canali as the products of advanced technology that captured the public imagination. Percival Lowell figures prominently in this story. He was the most prolific advocate of the advanced technology interpretation of the canali and after failing to gain an appointment at an established observatory, he used his wealth to construct one of his own dedicated to further examination of Mars. Lowell wrote numerous articles for the public describing his observations and promoting his picture of a Mars inhabited by intelligent beings. These articles appeared in leading literary magazines of the day (Lowell, 1895a; Lowell, 1895b; Lowell, 1895c; Lowell, 1895d). Among other things, Lowell set the stage for a generation of writers who have given us a vision of Mars teeming with life, often in violent conflict (Wells, 1898; Burroughs, 1912). Indeed, Lowell's Mars sounds so familiar to the modern ear in large part because of the later fictional works just cited. It requires a conscious effort to appreciate Lowell's original if scientifically dubious theories.

While his conclusions appear ludicrous today, his enthusiasm has surely motivated the scientific study of Mars. While the martian civilization hypothesis never gained widespread support in the scientific community, for much of the pre-spacecraft part of this century, a vegetation hypothesis for martian dark markings was widely entertained. The need for inorganic alternatives to such theories was demonstrated in the late 1950s (Kuiper, 1957). Using a color chart supplied by the Glidden Paint Company, Kuiper showed that the martian dark regions were not green as widely supposed, but gray or even the same red color as the rest of the planet only darker. That is one example of the considerable resources we have devoted to studying the possiblilty of life on what is in most respects a thoroughly inhospitable world. Despite frigid conditions, sterilizing UV radiation, and cosmic ray bombardment, scientists continue to make claims for either extant or fossil life
on Mars (McKay et al., 1996). None of these claims to date have received more than temporary acceptance or interest. In the final analysis, nanometer-scale blobs may be the latest version of thoats and beautiful princesses.

Not that I am immune. Mars has long been the planet I have most favored for study. The similarities to Earth are compelling, the prospect for colonization enticing. After the dizzying scramble from the Wright brothers to Neil Armstrong in just 63 years, prospects for a visit to Mars looked very good. For a variety of reasons, plans for continued manned exploration and colonization of the moon was abandoned. Also abandoned were plans for manned exploration of Mars, to have begun as early as 1984. Instead, here we stand now at the brink of the millenium and still human beings have not been to Mars. Instead, further exploration has been limited to robotic probes we have sent to Mars, and examination of bits of Mars that have come to us.

2.2 Rocks from Mars?

The notion that debris might be launched from a planet via impact is at least 60 years old. It was in 1936 that Nininger first thought of the possibility that tektites were lunar impact debris (Wilhelms, 1993), though it was many years before he published this hypothesis (Nininger, 1943a; Nininger, 1943b). Tektites are not a good analogy for the SNCs, since the SNCs are intact (solid) while the tektites are melted glass. The first explicit statement that material could be launched relatively intact from a planetary body appeared in 1963 when Shoemaker and colleagues mentioned the possibility in print (Shoemaker et al., 1963). The motivation for that work was the use of exchanged impact ejecta to correlate the geologic histories of various planets. This was presented in the same spirit as terrestrial geologists using rock samples to correlate the stratigraphic record between different regions of the Earth.

That material could be ejected from the moon was argued on the basis of
the distribution of secondary craters and laboratory experiments (Charters, 1960; Shoemaker et al., 1963) and was implicit in Urey's model for the origin of chondritic meteorites (Urey, 1959; Urey, 1967), though he frequently derided the "tektites from the moon" people (Urey, 1963; Wilhelms, 1993). Ejection of chondrites from lunar-sized bodies during the earliest stages of solar system formation was advocated even earlier (Urey, 1956). Regarding the proposal that meteorites could come from the moon, none of the authors included a detailed discussion of the relevant impact physics, that is, the specific launch mechanism. Documentation of the authors' ejection model, if one existed, is not available. Whatever the launch mechanism may have been, Shoemaker for one had clearly changed his mind on the subject by 1981 when he argued vigorously against the notion that intact, solid material could be ejected from a planet (Melosh, personal communication). Be that as it may, Shoemaker and colleagues had considered the possibility of martian meteorites in the 1963 paper. That work closed with a speculative paragraph about impact-ejected rocks from Mars, arguing that if debris can be ejected from the moon, then why not Mars as well? Coincidence, hunch, or prescience, the paper concludes with words that could have been published in Nature last week:

If some small amount of material escapes from Mars from time to time, it seems likely that at least some very small fraction of this material would ultimately collide with earth. Whether it could ever be recognized is difficult to say, but the possibility that such material could carry organic hitchhikers, however remote, may present a vexing question to those who are concerned with the origin of life.

Though decades-old, the quotation remains an apt appraisal the current consensus on the issue of interplanetary exchange of intact material.

It was in 1979 that the first very tentative proposals (that is really too strong a word) were made for a martian origin of the SNC meteorites (Wasson and Wetherill, 1979; Nyquist et al., 1979; Walker et al., 1979). Two years later, the
first strong argument for this notion was published (Wood and Ashwal, 1981). The authors began that paper with the famous quotation attributed to Sherlock Holmes: 
"...when you have excluded the impossible, whatever remains, however improbable, must be the truth." The quotation aptly characterizes the nature of the arguments presented for a martian origin for the SNCs. They were based on a large body of circumstantial evidence aided by the process of elimination. I will next summarize the observations and arguments presented there.

2.2.1 Why the SNCs are martian

The first of these observations regards the relative abundances of the rare earth elements (REEs), termed abundance patterns. The abundance patterns of the REEs in the SNCs are highly fractionated (Wood and Ashwal, 1981). This is representative of a planet that has been geologically active for much of its history. In addition, the abundance patterns provide information pertaining to the pressure conditions under which the parent magmas were extracted from the planetary mantle (McSween Jr., 1985; Taylor and McLennan, 1985; Henderson, 1986). The abundance of REEs is influenced by the minerals present in the source region, which in turn are stable only under certain ranges of pressure and temperature. Pressure is related to depth in a straightforward way, making it possible to quantify further the previous argument. A possibly simplistic interpretation of the SNC REE data would indicate the presence of garnet minerals in the magma source region, necessitating a Mars-sized or larger planet. That conclusion is model-dependent. A safer conclusion is that the SNCs formed on an evolved planet such as Earth, Venus, or Mars (Ma et al., 1981; Nakamura et al., 1982; Longhi, 1991).

In the SNC meteorites, the abundance ratio of the radioactive elements K and U are similar to that of the Earth though different from the moon and the eucrites (the eucrites are achondrite meteorites that probably originated on the asteroid Vesta). In addition, the oxidation state of the SNC parent body is similar
to that of the Earth's upper mantle. Both of these factors argue for a large parent body (Wood and Ashwal, 1981).

The relative abundances of the three stable isotopes of oxygen have proven an effective means of classifying meteorites (Clayton et al., 1976; Clayton, 1993). Analysis of oxygen isotopes in the SNC meteorites show that they shared a common oxygen reservoir (Clayton and Mayeda, 1983; McSween Jr., 1985; McSween Jr., 1994). Moreover, this reservoir is distinct from that shared by the Earth-moon system, arguing against the Earth as a possible parent body (Clayton and Mayeda, 1996). There have been efforts to find terrestrial material that does not fall on the terrestrial fractionation (TF) line, typically by sampling material thought to have been sheltered from tectonic recycling. These efforts to date have failed. Hence, the fact that a sample lies off the TF line is a strong argument that it is neither terrestrial nor lunar.

The ages and isotopic evolution of these meteorites were unique at that time—it was still two years before the first lunar meteorite was recognized (Mason, 1982). The initial Sr and Nd isotopes measured in the SNCs indicated that while these rocks came from the same planet, they came from different isotopic reservoirs on that planet, indicating a complex geologic history unlike what one would expect for an asteroidal parent body. The radiometric dates for the SNCs were all more than 3 Ga younger than all other known meteorites which, without exception, formed 4.5 Ga ago. At the time Wood and Ashwal were writing, the SNCs were thought to date to 1.3 Ga ago. Now we know that the shergottites are much younger still (McSween Jr., 1994). This was an important clue because there is a well-known qualitative relationship between the size of a planetary body and the length of time it can sustain volcanism from endogenous heat sources. Because of the favorable volume to surface area ratio, larger bodies take longer to cool and hence maintain geologic activity for a longer period of time. This simple idea was borne out by spacecraft observations of planetary bodies during the "Golden Age." Small bodies like the moon and Mercury have apparently been geologically dead
for most of solar system history—large parts of their surfaces record the vestiges of heavy bombardment (Wilhelms, 1984; Wilhelms, 1993). The youngest lava flows on the moon are estimated to be approximately 2.5 Ga old, based on the calibrated lunar cratering curve. Large bodies such as the Earth and Venus retain no record of the era of heavy bombardment, while Mars fits somewhere in between. This argues against lunar sized or smaller bodies as a viable SNC parent body.

The SNC meteorites show no signs of magnetism, meaning that apparently they had not formed in the presence of a magnetic field. This result is "in contrast to all other meteorite types." Since nearly all meteorites are thought to derive from the asteroids, a meteorite sharing a unique characteristic such as this lack of magnetism would logically require a different source. This is an argument that the source is not asteroidal (Wood and Ashwal, 1981).

The SNC meteorites show some signs of shock, consistent qualitatively with their being ejected from a large planet via impact. Moreover, it is important to note that these rocks were not breccias—they do not include bits of material reminiscent of other meteorite classes. Conversely, no SNC material has ever been found in brecciated meteorites, indicating that SNC material had not mixed with the source material for any other achondrite. Again this argues against an asteroidal origin (Wood and Ashwal, 1981).

Given this suite of information, Wood and Ashwal then eliminated the suspects one by one. Mercury was ruled out on the basis of its great distance from Earth (in terms of its depth in the sun’s gravitational well) and its apparent surface age. Venus was ruled out on the basis of its thick atmosphere and strong gravity. This turns out not to be as strong an objection as it was then, based on the probable formation mechanism of the impact-related parabolas (Arvidson et al., 1991; Campbell et al., 1992; Vervack Jr. and Melosh, 1992). An additional factor was that the high surface temperature should substantially alter the rocks in ways not observed for the SNCs. For example, the argon should be completely baked out of venusian surface rocks, resulting in a K-Ar age corresponding to the transfer
time of millions of years (Head and Swindle, 1995). The Earth and moon were eliminated on the basis of the oxygen isotope studies. The eucrite parent body was eliminated on the basis of size (asteroidal) and age of all other samples from that body. This left Mars as the only possible, however improbable, suspect.

2.2.2 The physicists respond

This met with considerable skepticism, and for good reason—there was no credible launch mechanism known at the time. Ejection from volcanic vents is not tenable for a martian origin, despite the impressive size of the Tharsis volcanoes. During a volcanic eruption, the speed of expelled solid material is constrained by the enthalpy of the expanding gases. By equating enthalpy with kinetic energy one obtains the relationship 

\[ v_{ej} = \sqrt{\frac{2h}{\mu}} \]

where \( h = C_pT = \frac{7}{2}RT \) is the enthalpy, \( \mu \) is the molar mass, \( R \) is the gas constant and \( T \) is temperature. For \( H_2O \) at 1200°C (earth upper mantle temperature) \( v_{ej} \approx 1 \) km/sec. This velocity is consistent with the maximum reported range (\( \approx 1 \) km) of volcanic bombs. The 100 km tall plumes on Io are consistent with the enthalpy of \( SO_2 \) (Keiffer, 1992). To get \( v_{ej} \approx 5 \) km/s (Mars’ escape velocity) for a water-rich mantle requires an eruption temperature of \( \approx 30000^\circ C \) (\( T \approx v^2 \)). To put it mildly, this temperature is much hotter than believed to obtain at the Earth's core and is not at all tenable for the martian mantle. The natural geologic cannon is just too feeble to launch rocks from Mars.

On the face of it, impact cratering is a more promising mechanism. Certainly an impact event transforms very large amounts of energy. It is easy to imagine some of this energy partitioned into ejecting fragments from the target body (Melosh, 1989). During an impact, target rocks are accelerated to very high velocities, velocities on the order of the impact velocity itself. Thus, accelerating material to martian escape velocity (5 km/sec) is not so difficult, given the median impact speed of asteroids with Mars of 7 - 12 km/sec (Bottke Jr. et al., 1994; Steel, 1996). The problem is reconciling the survival of solid material accelerated to that
speed with shock physics.

The relationship between the thermodynamic parameters in a shock were derived by P. H. Hugoniot from the conservation of mass, momentum and energy across the shock discontinuity. In the rest frame of the uncompressed material, these equations are:

\begin{align}
\rho (v_s - v_p) &= \rho_0 v_s \tag{2.1} \\
P - P_0 &= \rho_0 v_p v_s \tag{2.2} \\
E - E_0 &= (P + P_0) (V_0 - V) / 2 \tag{2.3}
\end{align}

where \( \rho_0, P_0, E_0 \) and \( V_0 \) are the density, pressure, specific internal energy, and specific volume (i.e., \( 1/\rho \)) of the uncompressed material (ahead of the shock), \( \rho, P, E \) and \( V \) are the equivalent parameters for the compressed material, \( v_s \) is the shock velocity, and \( v_p \) is the velocity of the material behind the shock (Melosh, 1989).

For impacts into realistic geologic materials, the second Hugoniot equation shows that a peak shock pressure of 100-150 GPa is required to accelerate material to Mars' escape velocity. As seen in Table 2.1 this pressure is far in excess of that required for melting—a rock accelerated to Mars' escape velocity should be completely melted if not reduced to vapor. These objections were raised vociferously at the LPSC meeting when Wood presented his conclusions. In particular, Gene Shoemaker raised this point, making his oft-quoted statement that if a rock is blasted from the surface of Mars, it is not going to be a rock anymore.

\subsection*{2.2.3 Loophole—interference with the free surface}

Of course, this objection ignored a volumetrically small, yet vitally important phenomenon present in an impact. The apparent minutia of impact physics (also seen in nuclear test explosions) that the above objection overlooks is the influence of the free surface on the structure of the shock wave. Shoemaker's (and many others—he was not alone in this opinion) objection was correct so far as the Hugoniot relations
Table 2.1: Petrographic shock indicators for various minerals (Melosh, 1989). Note the onset of melting at 45-65 GPa. This will provide an important constraint on the interpretation of my results.

<table>
<thead>
<tr>
<th>Material</th>
<th>Indicator</th>
<th>Pressure (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tonalite</td>
<td>shatter cones</td>
<td>2-6</td>
</tr>
<tr>
<td>quartz</td>
<td>planar elements and fractures</td>
<td>3-35</td>
</tr>
<tr>
<td></td>
<td>stishovite</td>
<td>15-40</td>
</tr>
<tr>
<td></td>
<td>coesite</td>
<td>30-50</td>
</tr>
<tr>
<td></td>
<td>melting</td>
<td>50-65(?)</td>
</tr>
<tr>
<td>plagioclase</td>
<td>planar elements</td>
<td>13-30</td>
</tr>
<tr>
<td></td>
<td>maskelynite</td>
<td>30-45</td>
</tr>
<tr>
<td></td>
<td>melting</td>
<td>45-65</td>
</tr>
<tr>
<td>olivine</td>
<td>planar elements and fractures</td>
<td>5-45</td>
</tr>
<tr>
<td></td>
<td>ringwoodite</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>recrystallization</td>
<td>45(?)-65(?)</td>
</tr>
<tr>
<td></td>
<td>melting</td>
<td>≥70</td>
</tr>
<tr>
<td>clinopyroxene</td>
<td>mechanical twinning</td>
<td>5-40(?)</td>
</tr>
<tr>
<td></td>
<td>majorite</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>planar elements</td>
<td>30(?)-45</td>
</tr>
<tr>
<td></td>
<td>melting</td>
<td>45(?)-65(?)</td>
</tr>
<tr>
<td>graphite</td>
<td>cubic diamond</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>hexagonal diamond</td>
<td>70-140</td>
</tr>
</tbody>
</table>

are concerned. The problem is that those equations are derived for shocks in a continuous medium. The effects of a boundary such as a free surface is not considered, and that makes all the difference. This is a point still misunderstood by many planetary scientists, even by some working in impact-related fields. In Figure 2.1 the isobars resulting from an impact are shown. The illustration is incorrect because the contours show that along the surface the pressure is simultaneously zero and very large—clearly a non-physical situation! What really happens is shown in Figure 2.2—the pressure contours bend in along the surface, reflecting the proximity of the zero-pressure boundary. Material in this interference zone can be spalled off the surface at planetary escape velocities without being shocked to vapor.

The spall mechanism takes its name from a phenomenon observed and described in the nuclear test literature (Eisler and Chilton, 1964; Eisler et al.,
Figure 2.1: The plotted isobars (thin, hemispherical lines) for peak shock pressure are incorrect. Along the free surface the pressure must be zero at all times. The indicated pressures are in the 1-50 GPa range—a non-physical situation to put it mildly. It is somewhat surprising how often this erroneous picture of the impact-induced shock wave is presented. The figure shown is from French (1998). Similar gaffes appear in Alvarez et al. (1995) and Taylor (1992), more than a decade after Melosh's first publication on spall in impacts (Melosh, 1984).

1966; Chilton et al., 1966; Viecelli, 1973). A nuclear test site is typically heavily instrumented with ground motion sensors. This allows detailed recording of the particle velocity as a function of time around the explosion site. The accelerometer data showed that near-surface layers of material parted from deeper material with which it had been in contact. This detached material is referred to as the spall. The physics is illustrated in Figure 2.3.

Immediately following detonation of a buried device, a compressive stress wave travels towards the surface. This wave reflects from the free surface as a tensional wave (Figure 2.4). In the spall zone, the tensional wave interferes with the still-oncoming compressional wave. The result is that the stress experienced by material in the spall zone is much reduced from what it otherwise would be, i.e., significantly less than that experienced by material beneath it. Precisely at the surface, the compressive stress is always zero. In addition, the particle velocity in the spall zone is enhanced by this interference effect. This phenomenon is known
Figure 2.2: The plotted isobars are now correct. Since the pressure at the free surface is zero at all times, the isobars bend in towards the impact site. The indicated pressures near the surface are much lower than in Figure 2.1, a phenomenon noted in numerical calculations by Ahrens and O'Keefe (1978). Note the physical state of target material denoted on the right side of the Figure. Material near the surface remains intact. Investigation of shockwave interference led to development of the spall model for ejection of meteorites from planetary bodies (Melosh, 1984).

as velocity doubling in which the particle velocity at the free surface is twice what it would be otherwise: \( v_{\text{efc}} = 2v_{\text{particle}} \) (Figure 2.5 through Figure 2.14). This is sometimes demonstrated in the lecture hall with a hammer and a properly prepared wooden dowel. A tap with the hammer sends a nearly one-dimensional stress wave along the dowel. If the end is weakly attached, it will fly from the end of the dowel with a velocity impressive compared to the gentleness of the hammer tap. This is relatively easy to simulate numerically. In impacts, a specially prepared surface is not required since the reflected wave causes a tensional stress greater than the rock's strength.

2.2.4 Stress waves in a 1-dimensional rod

As part of the familiarization process with SALE 2D hydrodynamic computer code used in this work, I simulated the propagation of a compressive stress wave down
Figure 2.3: The Figure is from Eisier and Chilton (1964). Part (a) is accelerometer data. The arrival of the compressive wave from the blast gives rise to the peak denoted $t_0$. The reflected tensional wave causes parting of formations at depth (b). The reflection of the tensional wave from this boundary causes the doubled peak ($t_1$). After rising to a maximum height (c), the spalled material falls back to earth, generating additional peaks in the accelerometer data ($t_2, t_3$).

A one-dimensional rod. I altered the BC and BCSET subroutines controlling the boundary conditions in order to impose a triangular pulse with a rise time of 10 msec and a decay time of 40 msec with a maximum amplitude of 100 m/sec in particle velocity. The rod was 125 m long by 10 m wide and composed of Westerly Granite with a resolution of 5 m/cell. I was able to track the propagation of the wave by noting when individual vertices either initiated or ceased motion. By noting time and distance, I was able to measure the wave speed as approximately 3500 m/sec. Calculating the sound speed from the input parameters gives 3400 m/sec. The measurement matches the calculation to within the quantization error.

As the wave approaches the end of the rod, it interacts with the free surface. The pressure gradient at the front of the wave steepens since the pressure at the tip must remain zero. The end of the rod exhibited velocity doubling. The wave reflected as a tensional wave propagating back up the rod. The tensional strains
Figure 2.4: The Figure is also from Eisler and Chilton (1964). It shows the direct compressive wave arriving at the free surface from the buried nuclear explosion. The wave reflects as a tensional wave which has a strength greater than the tensile strength of the geologic medium. The material parts were sufficient to initiate damage accumulation, which progresses up the rod directly behind the reflected wave.

Demonstrations of the spall phenomenon have been carried out under conditions more appropriate to meteorite impact studies (Gratz et al., 1993). In that work Gratz and colleagues report the results of a light-gas gun experiment conducted at the Lawrence Livermore National Laboratory. This involved shooting a penny-shaped (and sized) aluminum disc at a larger disc-shaped target of Westerly Granite (Figure 2.15). It was well-known at the time that in such experiments material is spalled off the back surface. However, Gratz was interested in material spalled in the opposite direction—back towards the gun barrel. The experiment was designed to capture this material. A capture fixture was attached to the barrel, with the idea that debris spalled from the target would penetrate it to some depth characterized by the fragment's size and velocity. Individual fragments could then be analyzed for petrographic features diagnostic of the peak shock pressure obtained.
by the fragment (see Table 2.1).

The results were unexpected. Material did spall off the front surface of the granite target, but in an amount and at such a speed that the capture fixture was destroyed. It was possible to determine only lower limits to the spall velocity, which was on the order of a kilometer per second. The petrographic analysis showed that the peak shock pressures obtained by the fragments were an order of magnitude less than that predicted by the Hugoniot equations.

The role of spall in impacts was investigated theoretically in a series of papers by Melosh (Melosh, 1984; Melosh, 1985; Melosh, 1987) and is discussed in more detail in Section 5.1. The implication of this model is that material near the surface can be accelerated to very high speeds without ever experiencing the high pressures one would expect in a continuous medium. The interference of the shock wave with the free surface provided a scenario by which material could be accelerated to escape velocity without being vaporized, or even highly shocked.

About the time Melosh presented his results on the spall mechanism for ejecting solid material from planets, the first lunar meteorites were discovered on Earth (Mason, 1982). This was direct evidence that solid material could be exchanged between planets by impact. The spall mechanism also explained the presence of meteorites recovered in the Apollo lunar soil samples. Landing a meteorite on Earth is easy to explain by allowing the atmosphere to slow the incoming bolide, shedding its energy in the form of ablation and radiation until the surviving core of material is slowed to terminal velocity—a speed low enough that the meteorite survives impact with the ground intact. On the moon this mechanism cannot work. Instead, one must recall that during an impact, two shock waves are generated, one in the target and one in the projectile. When the projectile shock wave reaches the "back," it interacts with the free surface there, accelerating solid, weakly-shocked material to high velocity with respect to the projectile but low velocity with respect to the target. This allows a portion of the meteorite to soft-land, even on an airless body.
After the Wood and Ashwal paper was published it was recognized that gases were trapped in the maskelinite (shocked plagioclase) phase of EETA79001 (Bogard and Johnson, 1983). Analysis of these gases showed that, after suitable correction for terrestrial contamination, they were a "dead ringer" for the atmospheric gases measured by the Viking landers on the surface of Mars (Becker and Pepin, 1984). The absolute abundances for 7 isotopes from 6 gases match over a range of 8 orders of magnitude. This provided the coup-de-grace so far as martian origin of the SNCs is concerned, and little mainstream opposition to the notion has been heard since. In 1994 a non-SNC martian meteorite was identified (Mittlefehldt, 1994). In the discovery paper, Mittlefeldt wrote that it was time to rename the SNC meteorites. Since he had identified as martian a meteorite that did not fit into the S,N,C classification, SNC was no longer an accurate name. He proposed that henceforth these meteorites be referred to as "martian," just as meteorites from the moon are referred to as "lunar." This naming convention I will follow for the balance of this work.

2.3 Hydrocode Modeling of Impacts

Hydrocodes work by solving exactly the finite-difference equations governing the material behavior under a variety of conditions including shock. The accuracy of the results are limited by knowledge of the material properties and the resolution of the finite difference approximations. Hydrocode modeling of impacts owes much to the nuclear testing program in both the United States and the former Soviet Union. Both nations developed hydrocodes to simulate their nuclear test data. Clearly it is advantageous to have some ability to predict the weapon's effects before executing the test. When the United States signed the Limited Test Ban treaty in 1963 it did so in part because of the confidence gained in the predictive power of such codes. Apparently, the thinking was that underground testing would provide sufficient data for code verification, eliminating the need for above-ground testing to determine the effects of interest. Since a nuclear explosion is the only human activity comparable
in violence to an impact event, it is only natural that the respective sciences of each phenomena borrow from each other. As an example, the code used in this study is a modified version of one developed at the Los Alamos National Laboratory to simulate nuclear explosions. There are a large number of hydrocodes in use today for a variety of purposes. This brief review will be restricted to those most relevant to the study of impact physics.

Hydrocodes have been used to simulate impacts for more than 25 years (O'Keefe and Ahrens, 1975). Early works appear to concentrate on simulating the development of the shock wave and its effect on target materials. This was certainly motivated by the need to understand the abundance of shock glass in the recently obtained lunar samples. The lunar samples reflected a terrain dominated by the effects of impact cratering. A quantitative understanding of shock effects at the appropriate size scales i.e., much larger than the laboratory, was required to evaluate the lunar samples in their proper context. Major improvements have been made through the years, and the effort by modelers to keep pace with new impact-related phenomena continues to the present day (Housen and Holsapple, 1999).

Eulerian hydrocodes were used to study impact ejecta in the late 1970s (O'Keefe and Ahrens, 1977; Ahrens and O'Keefe, 1978). The 1977 work simulated the impact of anorthositic and iron projectiles onto an anorthosite halfspace. That work determined the mass of ejecta (relative to the impactor mass) that achieved escape velocity as a function of impact speed for different target planets. Under certain conditions, the ejecta mass exceeds the impactor mass, meaning that the event is a net loss of mass for the planet—negative accretion or "de-cretion." That work quantified an idea articulated a decade earlier (Gault et al., 1963). The 1978 work continued the examination of ejecta from impacts, using the hydrocode model to study differences in ejecta implacement on the moon and Mercury. One of the peripheral results noted in that work was that the coolest ejecta came from the near-surface region near the impactor. This phenomenon was not examined in any
greater detail in that paper, but it clearly foreshadows the work on spall cited in the previous section. Had this point been followed through to its logical conclusion, it is possible that the existence of intact ejecta (meteorites) from planetary bodies such as the moon and Mars might have been predicted from a sound physical basis. At this time, the notion of a martian origin for some of the basaltic achondrites must have been discussed informally since it appeared in print a year later (Wasson and Wetherill, 1979). That reference is a chapter in Asteroids of the Space Science Series. Given the lead-time for some volumes of that series, it seems likely that by 1978 or even earlier, drafts containing the martian origin hypothesis were in existence, possibly being shared at conferences. Perhaps the modelers did not hear of it, or failed to recognize its importance and relevance to their own findings. In any case, it is clear that the geochemists arrived at the correct conclusion first and it was sometime later before the theoretical framework was erected.

Soon after publication of the spall model (Melosh, 1984; Melosh, 1985; Melosh, 1987), there was a study arguing that the basis for that model was incorrect (Holsapple and Choe, 1988). The Melosh spall model assumed that the interference zone arose from the finite rise time of the stress wave, given by $a/U$, where $U$ is the impact velocity and $a$ is the projectile diameter. The Holsapple and Choe (1988) model was based on a more sophisticated (alternating Lagrangian/Eulerian) code than used by Ahrens and O'Keefe. It was able to carry out the same calculation and keep precisely defined material boundaries. It is not clear that the Holsapple and Choe work was published other than in abstract form. It addresses the issue of shock wave interference that was simulataneously being studied in the defense community. This topic will be discussed in Chapter 5.

However, it was the Melosh model of stress wave interference that was used to successfully analyze the results of impact experiments (Polansky and Ahrens, 1990). In that work the authors conducted numerous impact experiments with various projectiles into San Marcos Gabbro. After each impact, the target was cut through the epicenter to obtain a good cross-section. Several distinct fracture
patterns characterized the target, including a prominent fracture that delineated the interference zone. Using the impact conditions in the Melosh model, Polansky and Ahrens noted good agreement between the measured fracture pattern and the theoretical interference zone.

Reliable hydrocode calculation of fragment sizes was not achieved until the 1990s (Melosh et al., 1992; Ryan, 1992; Asphaug, 1993). For the first time a hydrocode could simulate an impact and accurately predict the resulting fragment size statistics. This ability had eluded modelers at Los Alamos National Laboratory for some time and was widely regarded as an important breakthrough. Since then, Asphaug has continued to develop more sophisticated fracture models that rely on explicit (rather than statistically defined) flaws (Asphaug et al., 1996). The latest improvement in hydrocode modeling of fragmentation in impacts is developing a method of achieving high resolution in the spall zone, which will be described later. Much work remains to be done. There is still no satisfactory model of porosity, which is certainly a feature of many planetary materials. Developing such a model should prove of great importance in understanding ejection of meteorites from the moon, crater formation on asteroids, and possibly creation of asteroid families.

2.4 Interdisciplinary Conflict

It is the hallmark of a well-understood phenomenon that the interpretations and descriptions from the point of view of different scientific disciplines are congruent. The epistemological power of different fields of study provide their unique insights, yet no major discrepancies remain. A simple example is the age of the Earth: the value given by geochemists is in accord with astronomy, paleontology, stratigraphy, and evolution. Such relative harmony has not been the case with the martian clan meteorites. There are two major conflicts to be resolved.

The first is the great disparity (by one to several orders of magnitude)
between the number of impacts geochemistry requires to launch the martian meteorites and the number that can be accommodated by impact physics and what is understood of the cratering flux at Mars. A number of writers have either advocated or assumed a single crater as the source for all martian meteorites known at the time of publication (Wood and Ashwal, 1981; Nyquist, 1983; Bogard et al., 1984; Vickery and Melosh, 1987; Mouginis-Mark et al., 1992). This parsimony was driven in part by the 1) results from theoretical and one-dimensional numerical studies of spall in impacts and 2) ignorance of the results of orbital integration studies conducted much later. The theoretical and one-dimensional spall studies suggested that the crater size required to launch meteorites from Mars is rather large, on the order of 12 km. This in turn implies a small number of impact events based on current knowledge of the martian cratering flux (Vickery and Melosh, 1987; Warren, 1994). The simplest model of course is to require a single impact into the right terrain to explain the petrology, then rely on cosmic ray bombardment, in-space breakup, and luck in transit to explain the geochemical details. The high-resolution simulations I have conducted have greatly reduced (by about a factor of four) the size of crater required to eject the martian clan meteorites. Hence, my contribution to this debate is that while it is not as easy as some advocate (Bhaktivedanta Swami Prabhupada, 1985) to transport a rock from one planet to the next, it is much easier than previously thought. As I will describe later, my results resolve this first conflict.

The second conflict is between the observed age distribution of martian meteorites (geochemistry) and the inferred age distribution of martian terrain units (photogeology). Approximately two-thirds of the martian meteorites have crystallization ages on the order of 200 Ma. The ages of martian terrain units can be estimated from crater counts, though with rather large error bars. Terrain units of shergottite age appear to cover only 2% of Mars, based on Viking data (Tanaka et al., 1992). The most recent results from the Mars Global Surveyor mission improves this somewhat, but by not nearly enough (Alfred McEwen, personal communication). Why should most of the meteorites come from the smallest terrain units?
This discrepancy was also a factor in choosing a very-small-number-of-craters origin for martian meteorites discussed above. If the number of impacts is small, one can appeal to the statistics of small numbers. This trend continues to the present day, driving authors to advocate a minimum number of craters even at the expense of adhering closely to the geochemical data (Nyquist et al., 1998). To the extent that my simulations model realistic geologic materials and stratigraphy, I will present scenarios that resolve this conflict at a semi-quantitative level.

2.5 Conclusion

The martian origin for the SNC meteorites has been well-established though major discrepancies remain. These discrepancies are best illustrated when one examines the arguments over how many craters are required to account for the rocks in our collections. Many of the lower estimates have been driven by an overreliance on an insufficiently understood field: impact physics. The higher estimates have basically ignored the quantitative results of theoretical spall models, rightly judging that our understanding of geochemistry and celestial mechanics is more mature. Complicating the picture is the age-frequency distribution of samples vs. martian terrain units. By simulating impacts into complex terrains with high-resolution in the spall zone, I am able to provide significant insight into resolving both of these quandaries.
Figure 2.5: This is output of the velocity field from a SALE2D simulation of a one-dimensional stress wave in a 10 x 125 m rod. A triangular pulse with a peak particle velocity $v_{\text{peak}} = 100 \text{ m/sec}$ is imposed on the top boundary and the disturbance is allowed to propagate down the rod. By noting the time when the peak of the stress wave passes through any two cells, one can confirm that the simulated stress wave is traveling at the longitudinal sound speed (3400 m/sec) for the simulated material, Westerley Granite in this case (cf. Figure 2.7). The plot is scaled to the maximum particle velocity at each timestep, so one must examine the numerical scale to see that $v_{\text{max}}$ for material at the end of the rod is double that of $v_{\text{peak}}$ (cf. Figure 2.6).
Figure 2.6: Here the peak of the stress wave has reached the end of the rod and the particle velocity at the tip is twice that of the input wave $v_{peak}$.
Figure 2.7: Comparing this and the previous figure, one can see that the tail of the wave has moved approximately 30 m in 10 msec, or 3000 m/sec. This is the sound velocity one calculates for the material parameters used in this simulation (see Table 5.2).
Figure 2.8: These are isobars corresponding to Figure 2.5. Note that the tail is much longer than the front, as defined by the rise and decay times for the wave. The rise time is 10 msec, during which time the wave progresses \( \sim 30 \text{ m} \) at 3400 m/sec. Comparing this with the previous Figure, one can see that the width of the wave front is about 6 cells, or 30 m.
Figure 2.9: Here the stress wave is interfering with the free surface. Note that the wave front has steepened considerably and the maximum pressure in the simulation is somewhat reduced.
Figure 2.10: The stress wave has reflected as a tensile wave, the front of which is interfering with the oncoming tail of the compressive wave. Recall that at this time velocity doubling is occurring at the tip.
Figure 2.11: The tensile wave has caused fragmentation in the lower end of the rod. The damage contours indicate that the upper 3/4 of the rod is still intact. Note that the damaged zone lags just behind the tensile wave.
Figure 2.12: Contours for maximum pressure attained for each cell in the computation. The entire rod experiences the same maximum pressure except for the end. Approximately the last 15–20 m experiences reduced pressures due to near-surface interference. Note that the interference zone is half the width of the wave front.
Figure 2.13: Plot of fracture times. The damage initiated at the end of the rod and propagated towards the top, following the reflected tensile wave.
Figure 2.14: Computed fragment sizes in the rod. In the region traversed by the tensile wave the fragment size is uniformly small. Near the top, the tensile wave has not yet passed. At the bottom, near-surface interference effects have influenced the computed fragment size.
Figure 2.15: The figure is from Gratz et al. (1993). It depicts the experimental apparatus for recovery of spall fragments in a high-velocity collision. Examination of the recovered fragments revealed peak shock pressures of approximately 5 GPa, an order of magnitude weaker than that predicted from Equation 2.2. The recovery fixture was destroyed, so it possible to establish only a lower limit of 1 km/sec to the spall velocity.
CHAPTER 3

The Martian Clan Meteorites

Much has been deduced about martian petrographic evolution as a result of extensive geochemical analysis of the martian meteorites. This is the great advantage of having samples from Mars: it allows the full weight of modern scientific analysis to be brought to bear. The chief hinderence to a comprehensive understanding of these rocks is that martian meteorite provenance is not known in any detail, other than the obvious conclusion that these rocks were once near the surface of Mars. Be that as it may, an enormous amount of information has been gleaned from these samples, much of which is of great help in deducing the history of these rocks and how they came to Earth. It is the purpose of this chapter to review relevant data from martian meteorite geochemistry and recent advances in modeling the delivery of material from Mars to Earth. The implications this has for determining the number of impact events required to explain all the martian meteorite data will be discussed.

3.1 Review of Martian Meteorite Geochemistry

As implied above, the full suite of geochemical analysis has been applied to the martian clan meteorites. This has resulted in a number of datable events, the interpretation of which has been, and remains, subject to some controversy. The dated events correspond to core formation, crystallization of the rock, subsequent shock events, the launch event, and landing on Earth. In addition, cosmic ray exposure (CRE) data provides a clue about the timespan a sample has been in a
<table>
<thead>
<tr>
<th>Meteorite</th>
<th>Circumstance</th>
<th>mass (kg)</th>
<th>Meteoroid (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shergottites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) EET 79001</td>
<td>found, 1979, Antarctica</td>
<td>7.94</td>
<td>0.22</td>
</tr>
<tr>
<td>2) DAG 476</td>
<td>found, 1998, Libya</td>
<td>2.015</td>
<td>0.14</td>
</tr>
<tr>
<td>2a) DAG 489</td>
<td>found, 1997, Libya</td>
<td>2.146</td>
<td>0.14</td>
</tr>
<tr>
<td>Total paired</td>
<td></td>
<td>4.161</td>
<td>0.17</td>
</tr>
<tr>
<td>3) Shergotty</td>
<td>fell, 1865, India</td>
<td>4</td>
<td>0.17</td>
</tr>
<tr>
<td>4) Zagami</td>
<td>fell, 1962, Nigeria</td>
<td>18</td>
<td>0.28</td>
</tr>
<tr>
<td>5) QUE 94201</td>
<td>found, 1994, Antarctica</td>
<td>0.012</td>
<td>0.03</td>
</tr>
<tr>
<td>6) ALH 77005</td>
<td>found, 1977, Antarctica</td>
<td>0.483</td>
<td>0.08</td>
</tr>
<tr>
<td>7) LEW 88516</td>
<td>found, 1988, Antarctica</td>
<td>0.013</td>
<td>0.03</td>
</tr>
<tr>
<td>8) Y 793605</td>
<td>found, 1993, Antarctica</td>
<td>0.016</td>
<td>0.03</td>
</tr>
<tr>
<td>Nakhlites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9) Nakhla</td>
<td>fell, 1911, Egypt</td>
<td>40</td>
<td>0.37</td>
</tr>
<tr>
<td>10) Lafayette</td>
<td>found, 1931, Indiana</td>
<td>0.8</td>
<td>0.10</td>
</tr>
<tr>
<td>11) Governador Valaderes</td>
<td>found, 1958, Brazil</td>
<td>0.16</td>
<td>0.06</td>
</tr>
<tr>
<td>Chassignites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12) Chassigny</td>
<td>fell, 1815, France</td>
<td>4</td>
<td>0.17</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13) ALH 84001</td>
<td>found, 1984, Antarctica</td>
<td>1.931</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Table 3.1: Physical data for the martian clan meteorites. Approximately half the samples are found in Antarctica at a rate of about one every three or four years. The mass is the combined mass of all fragments from the named sample. The diameter of the implied meteoroid is calculated by assuming 50% ablation and a density of 3000 kg/m³. Data compiled from Warren (1994), Score and Lindstrom (1990), Satterwhite and Mason (1991), Score and Mason (1995), Kojima et al. (1997), and Grossman (1995). Note that the data in Warren for EET79001 are incorrect.

certain environment. In particular, $4\pi$ (CRE) ages indicate the length of time the sample was in space in a fragment meter-sized or smaller. Finally, the terrestrial age can be deduced for meteorite finds, i.e., those not observed to fall. Thus, one can attempt to construct a coherent history of the sample based on the geochemical data.
3.1.1 Shergottites

The shergottites are divided into two groups, each meriting a separate meteorite class, but which for historical reasons remain categorized together. These groups are the basaltic shergottites and the lherzolitic shergottites (McSween Jr., 1994).

The basaltic shergottites include Shergotty, Zagami, EETA79001, DAG476, and DAG489. These rocks do not represent liquid phases, but formed either in a near-surface magma chamber or as lava flows with entrained crystals. The initial Sr, Nd, and Pb isotopic composition of these rocks requires that Shergotty and Zagami formed in a magma chamber separate from EETA79001 (Jones, 1989). The more recently discovered meteorite QUE94201 probably best fits into the basaltic shergottite category. Martian atmospheric gases have been detected in EETA79001, Shergotty, Zagami, and QUE94201 (Bogard and Johnson, 1983; Becker and Pepin, 1986; Terribilini et al., 1998).

The lherzolitic shergottites are ALH77005, LEW88516 and Y-793605. These are unlike the basaltic shergottites in that they clearly formed in subsurface environments and thus are true cumulate rocks. These three rocks are so similar in their trace element chemistry that not only did they come from the same igneous layered intrusion, but probably from the same small region within that intrusion (Warren and Kallemeyn, 1997). Their radiogenic isotopes are consistent with their crystallization from the same magma as EETA79001 but distinct from that of Shergotty and Zagami (McSween Jr., 1994). Martian atmospheric gases have been detected in ALH77005, EETA79001, LEW88516, and Y-793605 (Becker and Pepin, 1984; Terribilini et al., 1998).

More should be stated about the last two samples Dar al Gani 476 and 489, discovered in the Libyan Sahara in 1997 and 1998. The samples have been available to investigators for just about one year as of this writing. Most of the data I have collected on these two stones is available in abstract form only. It is possible that this data will be revised in the future.
In general, the shergottites are the youngest known martian rocks. They come from on or near the surface. They all have experienced significant shock, with inferred peak pressures in the 30-50 GPa range (McSween Jr., 1985; Nagao et al., 1997).

3.1.2 Nakhlites

The three members of the nakhlite group are Nakhla, Lafayette, and Governador Valadares. Nakhla is noted for reportedly killing a dog when it fell in Egypt in 1911 (Graham et al., 1985). Lafayette is so named because it was found in a drawer at Purdue University in 1931. These rocks are all cumulate clinopyroxenites—true cumulate rocks that formed in a magma chamber rather than in a lava flow. At some point in their history they must have been uplifted to or near the surface in order to have been launched by a meteorite impact. Their Sr and Nd isotopic compositions are distinct from the shergottites and require a different parent magma (McSween Jr., 1994). All three nakhlites might have been comagmatic. Atmospheric gases detected in these meteorites have not been shock implanted, but rather incorporated either in secondary minerals or sediments cycled into the magma. No convincing shock features have been detected.

3.1.3 Chassigny

Chassigny is a dunite, an olivine-rich cumulate rock. This is another true cumulate rock that formed at depth, then was later brought to the surface by some unknown mechanism to facilitate launch from Mars. The REE pattern for this sample is sufficiently different from the nakhlites that they cannot be comagmatic. The initial Sr isotopic ratios are identical however. This leaves open the possibility that the shergottites, nakhlites, and Chassigny came from very different localities on Mars. Martian atmospheric gases have been detected in Chassigny (Terribilini et al., 1998). Chassigny has been shocked to 35 GPa (Langenhorst and Greshake, 1999).
3.1.4 ALH84001

This orthopyroxenite was originally misclassified as a diogenite before it was shown (Mittlefehldt, 1994) that it belongs to the martian clan meteorites. This rock appears to be a unique member of the martian meteorites. It has a truly ancient age of approximately 4.5 Ga and records several shock events (Treiman, 1995a; Treiman, 1998). The great age of this stone indicates it came from terrain surviving from the era of heavy bombardment, the only martian meteorite to do so. No Earth rock is so ancient. Martian atmospheric gases have been detected in ALH84001 (Swindle et al., 1995). This sample gained notoriety with the announcement of possible evidence for life on early Mars (McKay et al., 1996). The strength of the ejection shock is difficult to determine, given the complex shock history of this rock. Briefly, ALH84001 has sustained shocks of ~75 GPa, ~30 GPa, then ~60 GPa prior to the launch event, which generated a weaker shock (Treiman, 1998).

3.2 Martian Meteorite Ages

A number of ages relating to different events in the history of each martian meteorite can be deduced by various geochemical means. The age data available for these stones derive from the analysis of radionuclides or nuclear spallation products. The datable events range from core formation to ejection from the surface to terrestrial residence. Crystallization ages are generally derived from analyses of the Sm-Nd, Rb-Sr and U-Th-Pb systems. The K-Ar system less useful for this purpose because it is relatively easy to reset (from the shock of a nearby impact, for example). Shock ages are very difficult to determine. Published "shock ages" for various shergottites range from 180 Ma to zero years ago! As one might expect, the K–Ar system has been utilized here. CRE ages are determined from measuring the abundance of nuclear spallation products in the sample. Under cosmic ray bombardment nuclides split into measurable daughter products at a predictable rate. The abundance of the stable spall products indicates how long the sample has been exposed to the
space environment. The abundance of radioactive spall products indicates how long the sample has been protected by the Earth's atmosphere, i.e., its terrestrial age. Finally, martian crater statistics, combined with lunar crater and sample age data, provides some constraints on the age of martian terrain units that are ultimately the source of these rocks.

So far as the sample analysis is concerned, of most interest for my work are the crystallization and CRE ages. These tell how long ago the rock formed on Mars, which can be correlated with estimated ages for martian terrain units. This in turn allows a comparison between the age-frequency distributions of the meteorites and the surface of Mars. The CRE ages indicate when a sample was exposed to \( 4\pi \) cosmic ray bombardment, i.e., how long it was in space as part of a rock about 1 meter or less in diameter. This is not necessarily the ejection age, since in principle a rock large enough to shelter the interior from cosmic ray bombardment could be ejected then later broken into meter-sized pieces. In addition, there could be a non-negligible terrestrial age. In the case of martian meteorites, however, I will argue that CRE ages (plus terrestrial ages where appropriate) are most likely ejection ages. Ejection ages can be used (along with petrologic analysis) to constrain the number of impact events, which can then be compared to what is known about the impactor flux at Mars.

3.2.1 Crystallization ages

The crystallization ages for the shergottites are a subject of continuing controversy, but most likely ranges from 150–300 Ma (McSween Jr., 1994). The Sm–Nd whole rock data fall along a line (originally, and still occasionally interpreted as an isochron) corresponding to a 1.3 Ga model age (McSween Jr., 1985). Rb–Sr data for mineral separates gives isochrons for the basalts of 180 ± 20 Ma and U–Th–Pb gives a nearly identical 190 ± 30 Ma age. Sm–Nd work gives an older date, approximately 360 Ma, though this date probably results from mixing between different mantle
and crustal reservoirs. The 180 Ma event was initially interpreted as a shock, possibly the shock ejecting the rocks from the martian surface. That these meteorites would survive in space for 180 Ma is highly questionable when one considers that the erosion timescale for chondrites is on the order of 30–60 Ma (Marti and Graf, 1992; Bottke Jr. et al., 1994).

Jones (1986) argued compellingly that the 180 Ma event is in fact crystallization. He arrived at this conclusion by considering the two lithologies preserved in EETA79001. The initial Sr isotope ratios for the two lithologies are distinct. If the 180 Ma event is a shock, this distinctness is difficult to explain. How could a shock mobilize Sr sufficiently to reset the Rb-Sr chronometer while preserving the initial Sr ratios? This line of reasoning was supported by shock experiments that failed to reset Rb-Sr (Nyquist et al., 1987). Later, it was possible to obtain a date from shock glass in ALHA77005 of 15 ± 15 Ma (Jagoutz, 1989). This strongly argued that the shock event occurred long after the 180 Ma event, and that 180 Ma probably represents the crystallization age.

There is widespread agreement that the crystallization age for Chassigny and the nakhlites is 1.3 Ga (McSween Jr., 1994). This date is based on the Rb-Sr, Sm-Nd, U-Pb, and Ar-Ar systems. In contrast to the rest of the martian meteorites, ALHA84001 has been dated to approximately 4.5 Ga, though analysis of this rock is complicated by at least two, perhaps as many as five shock events (Treiman, 1995a; Treiman, 1998). As a consequence, the quoted crystallization age has comparatively poor precision, but it certainly dates to the era of heavy bombardment or earlier. Crystallization age data and related information for individual martian meteorites is given in Table 3.2.

### 3.2.2 Crater ages

There are no absolute ages for any known location on Mars. All martian surface ages result from a rather long chain of inferences based on the observed crater
populations on Mars, assumptions about weathering on Mars, the observed crater population on the moon, the derived time-dependent lunar cratering flux, and the extrapolation of that flux from the moon to Mars. The best determined of these are the observed crater populations. There is an established absolute age system for the moon, but its best precision is only for the era of heavy bombardment (prior to 3.5 Ga). For more recent times, which is more relevant to interpreting the martian crater population and martian meteorite provenance, the lunar crater-age data are not well-constrained.

First of all, the youngest lunar dates derive from analyses of impact ejecta. The crater responsible for the particular ejecta studied can be fit in the crater population, tying down a date for craters of that appearance and surface density. The problem here is one of uniqueness—the crater responsible for the dated ejecta material may not be correctly identified. In the case of Tycho, the ejecta material was collected thousands of kilometers from the crater (Wilhems, 1984; Wilhelms, 1993). Obviously, if the source crater is misidentified, the resultant crater-flux curve is not reliable.

Secondly, the cratering flux curve is rather flat after heavy bombardment, meaning a small difference in measured crater density leads to a large difference in derived crater age. This is exacerbated by the fact that the precision of crater density measurements is poor when compared with radiometric dates. An examination of Figure 3.1 clarifies this particular shortcoming.

To determine an absolute age for martian terrain units, one must apply the lunar-derived impactor flux curve to Mars. This is fraught with difficulty since receding from the sun should lessen the number of impactors while approaching the asteroid belt should increase the number of impactors. The most popular estimate at present is Hartmann (1999). He prefers using a value 1.6 times the lunar flux for Mars (Basaltic Volcanism Study Project, 1981).

With all these caveats in mind, it is possible to estimate roughly the area
of Mars covered by terrain of any given age. Approximately 16% of Mars appears to be of Amazonian age, thought to correspond to 1.3 Ga old terrain and younger. Only 2% is thought to consist of late Amazonian, corresponding to 180 Ma old rocks (Tanaka et al., 1992). However, 62% of the martian meteorites come from 180 Ma old terrain. This is the second fundamental mismatch in the martian meteorite story discussed in Chapter 2, here between geochemistry and cratering studies.

To the degree that they are known, the values given above for martian surface area covered by terrains of a given age is almost certainly incorrect. Since 1992, the Mars Global Surveyor (MGS) has been launched to Mars and is actively acquiring data as this is being written. It has been reported to me that the 2% figure above is too small by a factor of about 2 or even a factor of ten if there are many thin flows (Alfred McEwen, personal communication). It would follow that the 16% figure is also too small by some unknown amount. In addition, the high resolution images from MGS show terrains that are being exhumed (McEwen, personal communication). This means that there are young terrains on Mars that were exposed to cratering for some (unknown) period of time, covered and shielded from small impacts for some other (unknown) period of time, and finally re-exposed some (unknown) time ago. One of the consequences of this is that there exist on Mars some terrains that are older than their crater ages would suggest. On the other hand, MGS is revealing terrains with such low density crater populations that their formal crater age is on the order of millions of years (Strom, personal communication), suggesting that Mars is still geologically active. In counterpoint to this discussion is the sample ALHA84001, whose very existence shows that there are still very ancient rocks on the martian surface.

This still leaves a profound mismatch between the age distribution of martian meteorites vs. the martian surface. It has been suggested that older martian rocks are too weak to survive acceleration to escape velocity, and hence young rocks should dominate the meteorite samples (Jones, 1989; Warren, 1994). This argument was not convincing to some (Mouginis-Mark et al., 1992) because of the observation
at the Viking landing sites of large numbers of boulders. This assumes of course that the boulders really are "strong." My results will show that the observation of Mouginis-Mark et al. is probably irrelevant. The nature of the near-surface material has a profound effect on the shockwave and hence spall velocity. The strength of the rocks may not matter if the underlying material reduces the peak spall velocities to less than Mars' escape velocity.

In summation, the 13 martian samples primarily originated on relatively young terrain. Nearly two-thirds of them come from terrain on the order of 180 Ma old, which corresponds to the Jurassic era on Earth. If the source craters for these meteorites are randomly distributed across the surface of Mars, these ages would indicate that the martian surface is quite young—far too young to be reconciled with the crater populations observed there. I infer from this (as have others) that the age distribution of the martian meteorites is providing information about the physical characteristics of the martian surface.

3.2.3 Cosmic ray exposure ages

Cosmic rays are high-energy particles present throughout interplanetary space. Most of these particles originate from outside the solar system and have a mean energy of about 10 GeV (Wasson, 1985). These cosmic rays are sufficiently energetic to induce nuclear spall reactions to a depth of more than a meter. The nuclear spall products can be measured, and when compared against meteorite composition a model exposure age can be determined (Wasson, 1985). Cosmic ray exposure dating is a means of determining how long a particular sample has been subject to cosmic ray bombardment, i.e., how long a sample has been within a meter or so of the surface of an airless or nearly airless body. The martian atmosphere is too tenuous to significantly shield the surface from cosmic ray bombardment, and hence for these purposes Mars can be considered airless.
<table>
<thead>
<tr>
<th>Meteorite</th>
<th>Type</th>
<th>Age (Ma)</th>
<th>Pressure (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shergottites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EET 79001</td>
<td>basalt</td>
<td>?</td>
<td>30-43</td>
</tr>
<tr>
<td>DAG 476</td>
<td>basalt</td>
<td>400-800?</td>
<td>40-50</td>
</tr>
<tr>
<td>DAG 489</td>
<td>basalt</td>
<td>400-800?</td>
<td></td>
</tr>
<tr>
<td>Shergotty</td>
<td>basalt</td>
<td>180</td>
<td>30-43</td>
</tr>
<tr>
<td>Zagami</td>
<td>basalt</td>
<td>180</td>
<td>30-43</td>
</tr>
<tr>
<td>QUE 94201</td>
<td>basaltic gabbro</td>
<td>330?</td>
<td>30-43</td>
</tr>
<tr>
<td>ALH 77005</td>
<td>lherzolite</td>
<td>187±12</td>
<td>30-43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>154±6</td>
<td></td>
</tr>
<tr>
<td>LEW 88516</td>
<td>lherzolitic gabbro</td>
<td>?</td>
<td>30-43</td>
</tr>
<tr>
<td>Y 793605</td>
<td>lherzolite</td>
<td>210±62?</td>
<td>30-50</td>
</tr>
<tr>
<td>Nakhlites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakhla</td>
<td>clinopyroxenite</td>
<td>1300</td>
<td>undetectable</td>
</tr>
<tr>
<td>Lafayette</td>
<td>olivine clinopyroxenite</td>
<td>1300</td>
<td>undetectable</td>
</tr>
<tr>
<td>Governador Valaderes</td>
<td>clinopyroxenite</td>
<td>1300</td>
<td>undetectable</td>
</tr>
<tr>
<td>Chassignites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chassigny</td>
<td>dunite</td>
<td>1300</td>
<td>35</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALH 84001</td>
<td>orthopyroxenite</td>
<td>4500</td>
<td>complex</td>
</tr>
</tbody>
</table>

Table 3.2: Selected properties of the known martian meteorites. These are arranged in reverse order of inferred ejection age as shown in Table 3.3 below. Age refers to crystallization age and pressure to peak shock pressure. Note the young crystallization ages for all dated samples except ALH84001. In comparison, all other meteorites date to 4.580 Ma ago. Youngest lunar rocks are estimated to be approximately 2000 Ma old, based on calibrated crater density data. Data compiled from McSween Jr. (1994), Nagao et al. (1997), Jagoutz et al. (1999), Langenhorst and Greshake (1999), and Nyquist (personal communication).

It is possible to determine the exposure geometry for a given sample. Cosmic ray exposure is either $2\pi$ or $4\pi$, depending on whether the sample is on the surface of a large body e.g. the moon or part of a meter-sized object in space. While $2\pi$ (as well as $4\pi$) cosmic ray exposure has been demonstrated for lunar meteorites, the martian meteorites show only $4\pi$ exposure, indicating an origin 2 meters or more beneath the surface. This depth is still well within the interference zone of an impact. Note that the shielding depth applies to fragments in space as well. The lack of $2\pi$ exposure indicates that if some of these samples were part of larger
fragments that broke up in space, the fragments were at least 4 meters in diameter and that the sample came from the core. The CRE data is very useful, and when combined with the additional geochemical data described earlier provides a good constraint on the total number of impacts required to eject these rocks from the surface of Mars.

The CRE data are summarized in Figure 3.2 and in Table 3.3 (Eugster et al., 1997; Terribilini et al., 1998). Note that the Figure reports ejection ages. The data for Dar al Gani 476 and Y 793605 are from Ludolf Shultz (personal communication). If each CRE age bin corresponds to an ejection age (as opposed to breakup in space) then it is clear that 6 or 7 impacts are required to account for the martian meteorites. Strong source crater pairing reduces this number of course, and this scenario is favored by Nyquist et al. (1998). They were largely driven by the assumption, which my work shows to be erroneous, that the required impactor size must be very large (rare) and is independent of the target composition. The older scenarios favoring a single impact (Vickery and Melosh, 1987; Mouginis-Mark et al., 1992) have been discredited both on the geochemical grounds described above (McSween Jr., 1994; Treiman, 1995b) and the latest results from orbital evolution models described below.

3.3 Delivery of Rocks from Mars to Earth

Exchange of intact material between planets was first suggested in print nearly 40 years ago (Shoemaker et al., 1963). Explicit reference is made in that paper to ejecta from Mars. Contemporaneous experimental work appeared to support the notion that solid material could be ejected from impact craters and launched into space (Gault et al., 1963), at least from the moon. Harold Urey had argued for a long time that the moon was the source of the chondritic meteorites (Urey, 1959; Urey, 1967), which logically implies that there must exist some mechanism to liberate this material. By the time Wood and Ashwal presented their arguments in favor of a
Table 3.3: Cosmic ray exposure and terrestrial ages for the martian meteorites. It is assumed that CRE commenced upon ejection, hence the ejection age $T_{ej}$ is the sum $T_{cre} + T_{ter}$. The basaltic shergottites and lherzolitic shergottites form distinct groups, consistent with their very different petrology. Chassigny, though different petrologically from the nakhlite, cannot be distinguished from them on the basis of CRE data. EET79001, the Dar al Gani stones, and ALH84001 are distinct from each other and the rest on the basis of ejection age. Data compiled from Eugster et al. (1997), Terribilini et al. (1998), Folco et al. (1999), and Nishiizumi et al. (1999).
No provision was made in that work for the ejection of larger fragments. A later numerical analysis (Ahrens and O'Keefe, 1978) of the distribution of ejecta from impacts noted that the coolest ejecta came from the vicinity of the free surface. The implications of this idea with respect to lunar or martian meteorites were not discussed.

3.4 Delivery Timescales

Setting aside the issue of launch mechanism for the moment, the transport of small objects through the inner solar system has been studied theoretically for quite some time, dating back to the early work of Öpik in the 1950s. The first published numerical models of meteorite transport date back 35 years (Arnold, 1965). Even the more restricted problem of Mars-to-Earth delivery has been studied for nearly 20 years (Wetherill, 1984). This is an area of ongoing research where significant advances have been made quite recently (Wisdom and Holman, 1991; Levison and Duncan, 1994). Owing in large part to the development of more efficient algorithms and ever faster computers, it is now possible to integrate directly the equations of motion for small bodies in the solar system under the gravitational influences of all the planets from Venus to Neptune (Gladman et al., 1996; Gladman, 1997). Gladman in particular has claimed that there are no longer any further improvements that can be made in terms of precision or completeness in addressing this problem. While one is often suspicious of such claims, especially when unmodeled, non-gravitational influences such as the Yarkovsky effect (Peterson, 1976; Rubincam, 1995; Farinella et al., 1998) are known, it is undeniable that modern techniques have supplanted earlier work in many cases. It even seems likely these new results will prove robust.
3.4.1 Ōpek methods

The first estimates for meteorite delivery time were made by Arnold (1965) using numerical codes based on the work of Ōpek (1951). Though Ōpek never wrote such codes himself, these are invariably referred to as "Ōpek codes" or "Ōpek–Arnold codes." In his 1951 paper, Ōpek derived a simple formula for calculating collision probabilities between two objects orbiting the sun. The inverse of the collision probability is the collision timescale. With this insight, one can write a code that can track the orbital evolution of a small body as it careens throughout the solar system. As originally conceived, these codes worked on the assumption that the orbits of small bodies were altered only by a close gravitational encounter with a planet. If the mean time between planetary encounters is long compared to the evolution of the planet's orbital parameters, then those parameters can be chosen randomly for each encounter. The orbital evolution of a population of small bodies can be tracked to produce a statistically significant result.

This method was used to constrain the origin of the meteorites (Arnold, 1965). Recall that at that time, Urey still championed a lunar origin for the chondrites. The Ōpek method was used to distinguish the orbital evolution of lunar ejecta and ejecta from the inner edge of the asteroid belt. These codes and their evolutionary cousins were later used to simulate the delivery of material to the Earth from the rest of the solar system (Wetherill, 1974; Wetherill, 1984; Melosh and Tonks, 1993; Bottke Jr. et al., 1996).

There are many difficulties in using the orbital perturbation approach, mostly arising from the fact that no attempt is made to calculate the actual trajectory of the test particle. For example, the Ōpek formulation assumed that planetary encounters would be from random directions. While this assumption usually holds, it is clearly not the case for lunar ejecta interaction with the Earth and may not be valid for planetary ejecta interacting with its planet of origin. Probably the most important parameter for delivery of martian material is resonances. It is possible
to account for resonances in an Ōpik–Arnold code, but this is done after-the-fact. It is not possible to know a priori which resonances will be important for a given problem. Avoiding these difficulties exist requires the use of other methods. Fortunately, fast, efficient methods for direct integration of the orbits of small bodies have been developed (Wisdom and Holman, 1991) which with modern computers can be used to study many problems of interest (Levison and Duncan, 1994) including delivery of the martian clan meteorites (Gladman et al., 1996; Gladman, 1997).

### 3.4.2 Direct numerical integration

About 10 years ago, a breakthrough of sorts was made in numerical codes for directly integrating the orbits of small bodies in the solar system (Wisdom and Holman, 1991). By integrating the orbits, the actual trajectory of the particle is calculated throughout the time period of interest, typically hundreds of millions of years. This had proved prohibitively expensive in computation time, hence the reliance on Ōpik codes. The development of symplectic integrators however greatly increased the computational efficiency to the point that it was possible to explore problems of interest in a reasonable amount of time. The high efficiency of these codes is accomplished by dividing the Hamiltonian of motion into two fully-integrable parts (Levison and Duncan, 1994) \( H = H_k + H_i \) where \( H_k \) is the sun-centered Keplerian orbit and \( H_i \) is the interactions between the test particle and the planets. If \( H_i \) is small, this formulation is rigorous. When \( H_i \) is large (during a planetary encounter), \( H_k \) is redefined to be planetocentric. This concentrates the available computational power on those parts of the problem most requiring it. Since gravitational interactions are calculated at each step, the effects of resonances are accounted for automatically. The chief drawback appears to be the inability to test all aspects of the code rigorously against analytical calculations.

The results of both approaches have been compared side by side (Dones
et al., 1999). This comparison is of particular relevance to the issue of martian meteorite delivery because the delivery efficiency predicted by the different methods differs by an order of magnitude (Wetherill, 1984; Melosh and Tonks, 1993; Gladman et al., 1996; Gladman, 1997). As can be seen in Table 3.4, the direct orbital integrations shows that Mars ejecta is delivered to the Earth much more quickly than the Ōpik approach calculates. This is due to the presence of a previously unsuspected combination of resonances operating in near-Mars space that increase the eccentricities of the small bodies. The orbits quickly become Earth-crossing. The resonance is so strong that the eccentricities are driven towards unity in a rather short timescale, at which point the orbits become sun grazing and martian material is lost. The most common fate for Mars ejecta is to fall into the sun.

<table>
<thead>
<tr>
<th>Fate</th>
<th>Integration</th>
<th>Ōpik</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 Ma</td>
<td>10 Ma</td>
</tr>
<tr>
<td>impact Venus</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>impact Earth</td>
<td>0</td>
<td>4.0</td>
</tr>
<tr>
<td>impact Mars</td>
<td>5.0</td>
<td>9.5</td>
</tr>
<tr>
<td>impact Sun</td>
<td>0</td>
<td>4.0</td>
</tr>
<tr>
<td>reach Jupiter</td>
<td>0</td>
<td>2.5</td>
</tr>
<tr>
<td>survivors</td>
<td>95.0</td>
<td>79.0</td>
</tr>
</tbody>
</table>

Table 3.4: Comparison of results from direct numerical integration and Ōpik orbital evolution calculations regarding the fate of martian impact ejecta. Figures are given as percentages. Of most relevance to this work is the time dependence for delivery of material to Earth. Under the numerical integrator, resonances work to increase the eccentricities to Earth-crossing values very quickly. Effects of these resonances are not included in the Ōpik calculation. Note also that a plurality of martian ejecta impacts the sun. Data from Dones et al. (1999).

3.5 Delivery of Martian Material to the Earth

Of the material that reaches Earth, about half does so in less than 10 Ma. This is in rough agreement with the CRE ages for martian meteorites, where the shergottites all have CRE ages much less than 10 Ma, while the nakhlites, Chassigny, and
ALHA84001 have larger CRE ages. This is assuming of course that the martian meteorites represent several impact events instead of a single large event. Given this assumption, the direct integration schemes are able to account for the CRE age distributions for the martian clan meteorites as well as the lunar meteorites. In the lunar case, the difference in model predictions is more extreme and it is obvious that the Ópik schemes are inadequate.

In his papers on the subject, Gladman concludes from the CRE data and his numerical integrations that martian meteorites are launched every million years on average. Accepting this conclusion as correct, it appears that the martian meteorites represent material from several (7-8) not a few (3-4) impact events. These impacts occur somewhat more frequently than the 10 Ma timescale for delivery of the ejecta to Earth. Thus, the martian meteorites represent a roughly constant rain of debris showering the Earth with bits of Mars. This is in contrast to the moon, where debris arrives at Earth with a timescale much shorter than the ejection event timescale.

3.5.1 Terrestrial ages and in-space breakup

The process of collecting meteorite samples introduces biases and constrains interpretations of the martian meteorite record. The terrestrial age of meteorites is typically very small compared to its transit time. In theory, some material from Mars is ejected directly into Earth-crossing orbits, implying that transit times on the order of months are possible. The required ejection speed ranges from 5.35 to 5.9 km/sec, depending on the launch direction and Mars' location in its orbit (Gladman, 1997). I will show simulations where some material is ejected at this speed or greater, in agreement with Gladman's conclusion on other grounds that such ejecta is fairly common. The known terrestrial ages for martian meteorites ranges from decades to 200 ka. If we consider only the Antarctic samples and assume that the greatest known terrestrial age is representative, then I calculate that roughly 1 fragment per million launched should be in the curated meteorite collections. The
museum efficiency for delivery of martian material can then be parameterized as follows

\[ E_m = E_{del} E_{coll} t_{earth} / t_{space} \]  

where \( E_m \) is the museum efficiency, \( i.e. \) the probability that a fragment ejected from Mars resides in a museum on Earth, \( E_{del} \) is the fraction of Mars ejecta that is eventually delivered to Earth in a delivery time \( t_{space} \) and \( t_{earth} \) is the terrestrial age of the sample. \( E_{coll} \) is a factor representing the areal fraction of Earth from which meteorites are collected with 100% efficiency. If we choose \( E_{del} = 0.05, E_{coll} = 0.001, t_{earth} = 200 \) ka, and \( t_{space} = 10 \) Ma, then \( E_m = 10^{-6} \). Since \( t_{earth} = 10 \) Ka may be more representative, it appears that each fragment of Mars in our collections represents \( 10^6 \) - \( 10^7 \) fragments we did not collect.

There are 4 known falls amongst the martian meteorites since 1815 (Table 3.1). Martian meteorites arrive at Earth approximately once a month (see below, and Section 6.1). The collection efficiency then is \( \sim 10^{-3} \) and the timescales of interest would indicate a museum efficiency of \( \sim 10^{-7} \) to \( \sim 10^{-8} \). Both analyses are limited by the small number of known samples (4 falls, 6 Antarctic finds).

This calculation argues strongly against shielding and in-space breakup as a significant process in the delivery of martian material to the Earth. A 10 meter diameter fragment can certainly shelter material from cosmic ray bombardment, but it is not large enough to accommodate \( 10^7 \) shergottite-sized fragments. In addition, one is left with the problem that not a single martian meteorite is known to possess a two-stage CRE history. A 10 m fragment would shelter no more than 10% of its volume from cosmic ray bombardment. These geochemical grounds, along with the above dynamical arguments, strongly favors a larger number of ejection events, implying modest source-crater pairing. This argues against in-space breakup as favored by various groups over the years (Vickery and Melosh, 1987; Mouginis-Mark et al., 1992; Nyquist et al., 1998). The results from my simulations to be discussed strongly favor multiple ejection events, consistent with the CRE data and analysis of meteorite collection efficiency.
This is significant because it pertains directly to the issue of how many impact events are required to launch the martian meteorites. In the extreme case (Vickery and Melosh, 1987) only a single impact is necessary, but this is untenable on the geochemical and dynamical grounds cited above. Even the less extreme position of one impact for each type of martian meteorite (Nyquist et al., 1998) requires three in-space breakup events for the shergottites, though none for the nakhlites, Chassigny, and ALHA84001. In addition, the breakup event must segregate lherzolitic material from the basaltic shergottites, as well as the surface rocks that would have recorded $2\pi$ cosmic ray exposure. In the absence of information about the size (frequency) of an impact required to launch material from Mars, one is driven towards the conclusion that a large number of impacts, perhaps as many as seven, are required to explain the geochemical data and remain consistent with the dynamical calculations. The dynamical argument in itself adds a constraint to the problem I am trying to solve. Remember that it is not enough to have an impact that ejects just one fragment to Mars' escape velocity, or even 20 fragments (so that one reaches Earth). In order to have any reasonable expectation that an impact capable of delivering material from Mars to Earth actually does so, one requires about $10^6-10^7$ fragments be accelerated to Mars' escape velocity. This constraint is in addition to the lack of two-stage CRE histories discussed above.

Consideration of terrestrial and CRE ages are useful in resolving the apparent discrepancy between the relative abundances of lunar and martian meteorites (Gladman et al., 1996; Gladman, 1997). The transit times of lunar meteorites are very much shorter than that of martian meteorites. Gladman argues that instead of a steady rain, we are seeing the ejecta from each lunar impact almost immediately afterwards. In other words, the timescale for ejection from the moon is large compared to the transit time. The situation is reversed for the martian meteorites. Gladman's estimate on dynamical grounds is that launch events take place on Mars every million years or so, possibly even more often (Gladman, 1997). As I will discuss later, my results show that this estimate was quite good, though somewhat pessimistic.
Given the data summarized in this Chapter, it is possible to estimate the flux of martian meteorites to the Earth (Gladman, 1997). There are approximately 20,000 known meteorites, of which order 10 are martian. The total meteorite flux has been estimated at 30,000 a year over the entire surface of the Earth (I. Halliday and Griffen, 1989). This implies a delivery rate of one martian meteorite per month on the Earth. This is probably an underestimate since the figure from Halliday et al. is conservative and pairing of Antarctic meteorites will increase the martian fraction.

3.6 Conclusion

While there are still many who argue that only a small number of launch events for the martian meteorites can be accomodated, it appears clear that in fact many impacts are required. At face value, the geochemistry suggests seven different impacts. What is known of the dynamics of transfer between planets is consistent with such a number. From a reading of the literature and discussion with the geochemists involved, it appears that a small-crater origin for the martian meteorites was avoided because of a lack of understanding of the details of spall in impacts. First of all, the size impactor required is much smaller than previously thought. This means that a model requiring on the order of ten impacts in the last 10 Ma is acceptable so far as impact physics and the inferred impactor population is concerned. Secondly, the nature of the target material has a profound influence on spall velocity and the volume of material ejected to escape speeds. Target properties can be invoked quantitatively to explain the age-frequency distribution of the martian meteorites with the singular exception of ALHA84001. It is the new found ability to conduct hydrocode simulations with high resolution in the spall zone for impacts into layered terrains that should settle this quandry.
Figure 3.1: Number density of lunar craters as a function of age determined from radiometric dating techniques. The crater density is determined from photogeology while the age data relies on samples. The most obvious feature of the curve is the non-uniformity of impact cratering through time. The earliest epoch (heaviest cratering) is referred to as "heavy bombardment." For most of lunar history (since about 3 Ga ago) the cratering flux has been rather light. The flatness of the curve and the inherent uncertainty in crater density data means that crater ages of terrains during this time are not well-calibrated. This curve provides the only basis for an absolute crater timescale on Mars. Figure from Wilhems (1984).
Figure 3.2: The histogram displays the number of martian meteorite samples of a given ejection age, based on CRE and terrestrial age data. Most samples have an ejection age of less than 4 Ma, all of which possess the youngest crystallization ages of approximately 180-300 Ma. The 1.3 Ga old samples and 4.5Ga old ALH84001 have older CRE ages. The CRE ages for Chassigny and the nakhlites overlap while those for the basaltic and Iherzolitic shergottites probably do not (see Table 3.3). The CRE data support as many as 7 ejection events for the martian clan meteorites (Eugster et al., 1997; Terribilini et al., 1998; Nishiizumi et al., 1999).
CHAPTER 4

The SALE 2D Hydrocode

As we have seen, analytic and one-dimensional numerical models proved incapable of producing solutions to the spall problem that were quantitatively consistent with the known geochemistry of the martian meteorites. Clearly, another line of attack was required. Assuming that spall is the responsible mechanism, what was needed was a numerical calculation in 2D with sufficient resolution to examine the stress wave interference with the free surface. Also required was multiple material capability to examine the role of layering in the target material, a geometry too complicated for an analytic treatment. The numerical model also needed a demonstrated capability to compute accurately fragment sizes resulting from the impact. Many of these requirements were met by the SALE 2D hydrocode (Amsden et al., 1980), a computer code designed to model fluid flow at all velocities. This code was significantly improved to incorporate material strength and fracture (Melosh et al., 1992) and is discussed in great detail in Ryan (1992) and Asphaug (1993).

This version of the code was tested extensively against all the laboratory fragmentation data available, including approximately 150 laboratory impacts conducted as part of Ryan's thesis work (Ryan, 1992). In that work, Ryan showed that the fragmentation model coupled to SALE2D accurately reproduced the fragment size distributions measured in the laboratory for a variety of impactor and target compositions. These experiments included impacts into targets comprised of weakly (1.6 MPa tensional strength) glued together pieces of gravel or fragments from an earlier impact test. In these cases the fragment size distribution closely resembled that of impacts into intact objects (Ryan, 1992). This is precisely the behavior one
expects in the dynamic fragmentation regime (see Subsection 4.1.1).

Asphaug used this version of SALE to examine the formation of Stickney on Phobos (Asphaug, 1993). Since then, Asphaug has continued to use evolved versions of SALE to study asteroid collisions (Asphaug et al., 1996; Asphaug, 1997). Here at Arizona, this hydrocode was improved by adding a multiple material capability in order to simulate the formation of dark splotches on Venus (Head and Melosh, 1995). Finally, after a great many blind alleys, a method was found to achieve high resolution in the spall zone (Head and Melosh, 1999a; Head and Melosh, 1999b). In this chapter I will briefly describe how the hydrocode works, concentrating on those parts most critical for the present study. I will then discuss the numerous upgrades to the code since it was last formally described (Melosh et al., 1992). Finally I discuss the process by which high resolution in the spall zone is achieved.

4.0.1 Why hydrocodes?

Hydrocodes numerically integrate the partial differential equations of mass, momentum, and energy conservation, together with an equation of state relating stress (pressure), strain (density), and temperature (internal energy). Since analytical solutions exist only for the most symmetric, idealized cases, the differential equations are rewritten as finite difference equations and solved numerically (Anderson Jr., 1987). By capturing the essential physics, numerical codes allow the investigation of problems poorly suited to experimental study. The chief limitations of this approach are the degree to which the properties of the materials of interest are known and the resolution of the finite difference equations.

Hypervelocity impacts of any significant size certainly qualify as poorly suited for experimental study. The scale of impacts considered in this work are about 500 MT. Obviously it is not desirable to conduct 500+ MT-scale experiments on terrestrial targets, no matter how well-instrumented. For many characteristics of crater formation, scaling laws have proven useful in calculating plausible impactor
properties for a given crater. While there are reservations about scaling events over 5 or 6 orders of magnitudes, these scaling laws are widely used with apparent success in addressing various problems pertaining to crater formation.

However, I must be able to calculate fragment sizes with a fair degree of confidence. In the past, fragmentation has been studied empirically using small-scale (~3-30 cm diameter targets) impacts. Although a great deal of data is available, it is unclear how to extrapolate these results to the size scales of interest. Various scaling laws exist to describe the laboratory data, however they give widely divergent results when extrapolated to objects 10–100 km in size (Fujiwara et al., 1989). Experiments are conducted over too small a size range to aid in selecting the correct scaling law. The method of solution employed by Melosh et al. (1992), Ryan (1992), and Asphaug (1993) was to develop a physical model for fragmentation that can be implemented in a fast numerical code. Once laboratory experiments can be duplicated reliably on the computer, one can have some confidence that the essential physics has been identified and included. Then one can use that computer program to simulate events beyond the reach of experimental investigation without relying upon extrapolation by some favored scaling law.

4.1 SALE

The SALE 2D hydrocode is approximately 3300 lines of FORTRAN designed to simulate fluid flow at all velocities. The original code is described in Amsden et al. (1980). Developed at the Los Alamos National Laboratory, this code was originally designed to study effects of nuclear weapons explosions in geologic materials, though the examples given in the 1980 report involve such mundane situations as supersonic flow through a pipe and collapsing dams. Given its pedigree, the hydrocode is well suited to the study of impacts since small impact events are similar in size and ferocity to nuclear explosions.

SALE works by establishing a two-dimensional grid composed of cells. The
cells are defined by vertices. The material parameters tracked during a run can be either cell-centered (pressure, internal energy, damage) or vertex-centered (position, velocity). During each computational timestep, the program iterates through the grid cell by cell, updating the material parameters and vertex velocities and positions by applying Newton's Second Law. This process is repeated for each timestep until the computational run is complete (Anderson Jr., 1987).

There are two modes of operation for the program, distinguished by how the computational grid is defined. In the first (Lagrangian mode) case, the grid points move along with the fluid flow. The benefit of this approach is that the material interfaces and free surface are always well-defined—a key requirement for studying shock wave interference phenomena. The chief drawback of this method is that cells can become highly distorted to the point that non-physical conditions obtain such as negative volumes and energies.

In the second (Eulerian mode) case, the grid is fixed in space and material fluxes through the cells. This approach avoids the problem of mesh entanglement but new problems arise. Material interfaces are not well-defined. Indeed, most of the time boundaries between different materials and the free surface itself are located somewhere between cell boundaries. Fluxing of even scalar parameters such as damage and fragment size (described below) is quite uncertain. In this work I used SALE in both modes in separate, well-defined steps.

4.1.1 Dynamic fracture

At low strain rates, fracture occurs by the process of gradual crack initiation and growth. The largest (weakest) flaw activates, growing in such a way as to relieve the applied stress. The material strength in this regime is a constant. At high strain rates (applicable to impacts) fracture occurs by the rapid growth of all available crack nuclei (Grady and Kipp, 1980). Fragmentation in this regime is controlled by the dynamic propagation of cracks and hence is known as dynamic fracture. The
applied stress field continues to grow during crack growth, leading ever stronger flaws to activate in addition to the weakest flaws. The transition strain rate between the static and dynamic regimes has been measured for concrete. The transition occurs over a broad range of $10^{-1}$ to $10^2$ /s, with a mean of about $10$ /s (Antoun and Rajendran, 1991). One of the consequences of this behavior is that the material strength is not constant in the dynamic regime, but increases with increasing strain rate. Laboratory measurements show that dynamic strength can exceed the static strength by an order of magnitude (Ryan, 1992). The growth of fractures alters the material properties, which then affects the propagating shock wave. The material response to the imposed shockwave must treat dynamic fracture correctly to produce an accurate solution for fragment size and other parameters of interest.

The fracture calculation is based on the Grady–Kipp theory of fragmentation (Grady and Kipp, 1980) utilizing the Weibull distribution of flaws (Weibull, 1939: Weibull, 1951: Irwin, 1958; Jaeger and Cook, 1969) It is assumed that rocks fail in tension. This assumption is based on the observation that the tensile strength of rocks is very small compared to the crushing strength. Flaws in the rock are activated by the release wave during an impact. As flaws are activated, cracks grow, relieving the applied stress field. Crack growth is characterized by a damage parameter, a scalar quantity that varies from zero (intact rock) to one (rubble). It is used to decrease the value of the elastic moduli in tension. When damage for a cell reaches unity, a fragment size for that cell is calculated and recorded.

4.1.2 Weibull distribution

The Weibull distribution assumes that in a given rock there is a large number of flaws which, when subjected to a given strain, will begin to fail. Flaws do not fail at the same strain. Instead, some flaws are small (strong) and only fail at large strains. Other flaws are large (weak) and fail at relatively small strains. The distribution is parameterized by Equation 4.1 below. It relates the number of activated flaws to
the material strain.

\[ N(\varepsilon) = k\varepsilon^m \] (4.1)

\( N \) is the cumulative number of flaws activated by a given strain \( \varepsilon \) and \( k \) and \( m \) are known as the Weibull parameters. Weibull parameter \( k \) has units of reciprocal volume and \( m \) is dimensionless (Weibull, 1939). Weibull's original formulation used stress instead of strain (Jaeger and Cook, 1969). The above formulation was used by Grady and Kipp and kept by Melosh et al. to avoid problems associated with strain softening. The numerical values for \( k \) and \( m \) are determined from tensile strength vs. strain rate experiments. This form of the distribution function has a wide array of applications (Weibull, 1951; Irwin, 1958), so Equation 4.1 is not closely tied to a particular physical model of material failure. Rather, it is a simple distribution that correctly reproduces the observed behavior of geologic materials in failure. In other words, the primary justification for its use in this context is that it gives satisfactory results.

Since the number density of activated flaws is proportional to strain, one might consider the strain required to activate a single (weakest) flaw in a given volume. This requires that \( N(\varepsilon)V \geq 1 \). The activation strain \( \varepsilon_{\text{min}} \) is derived then from

\[ k\varepsilon_{\text{min}} = 1/V \] (4.2)

This is known as the static threshold strain since flaws cannot be activated for strains less than \( \varepsilon_{\text{min}} \). Note the volume dependence. The larger the volume, the smaller the activation strain, meaning the weaker the material. This conceptual model is justified in Weibull's 1951 paper, where he uses a chain analogy for flaws in solid material. Consider a chain under tension. The probability of non-failure for a chain is identical to the probability of simultaneous non-failure of all the links in the chain. As links are added (volume of interest increases) the probability of non-failure decreases. Adding a link (increasing the volume) can only weaken the chain since no link can have infinite strength. Thus material strength is inversely proportional to volume.
This material behavior manifests itself in the current problem in defining the activation strain for the computational grid. As currently formulated, the activation strain depends on cell size using Equation 4.2. Larger cells are more likely to contain a weaker flaw than smaller cells. The issue of total volume of material in the global problem is ignored as in the past (Melosh et al., 1992; Ryan, 1992; Asphaug, 1993). In those works it was shown that the above formulation (in conjunction with Grady-Kipp fragmentation described below) worked very well. Hydrocode calculations were performed to simulate numerous laboratory collisions. The computation was able to reproduce with high-fidelity the fragment-size distribution obtained in the experiments. This lends some degree of confidence to the ability of the code to estimate the size of fragments ejected during an impact. On a historical note, one of the original motivations for characterizing flaws in rock was to predict the fragment sizes in quarry blasts. There is an optimal size into which to break quarry rocks. Fragment sizes that are too small (powder, for example) may be worthless while fragments that are too large means the rocks must be further reduced by manual labor, which is both expensive and dangerous (Ivanov, personal communication).

4.1.3 Grady-Kipp fragmentation model

Once flaws are activated they are assumed to progress at constant speed known as the crack velocity \( c_g \), which is typically about 0.4 of the longitudinal wave velocity \( c_l \) (Irwin, 1958; Melosh et al., 1992) of the material given by

\[
\sqrt{(K + (4/3)\mu)/\rho}
\]

where \( K \) is the bulk modulus, \( \mu \) is the shear modulus and \( \rho \) is the density. As cracks progress, they begin to intersect. Eventually they coalesce, dividing the formerly intact rock into a number of fragments.

As the cracks propagate, they relieve the stress in a volume \((4/3)\pi a^3\) where \( a \) is the crack half-length. Flaws inside this volume are sheltered from the applied stress field. This influences the size of fragments generated by the activated flaws.
and explains the observed strain-rate dependence of strength. If the stress field is applied very quickly (high strain rate), the volume of stress-relieved material grows slowly in comparison. This means that strains achieved in the unrelieved portion of the rock are larger than they would otherwise be, meaning that more (stronger) flaws are activated. This leads to a larger number of fractures which when they coalesce generates a larger number of smaller fragments. Since it is the stronger flaws that are activated, the rock behaves as if it is stronger than at lower strain rates. At the lower strain rates, the volume sheltered from the stress field grows rapidly and fewer flaws are activated. This leads to a smaller number of larger fragments and an apparently weaker rock, since it is the larger, weaker flaws that are activated.

As part of her thesis, Ryan compared the results of the fragment model predictions of fragment size distributions with results obtained in collisions under laboratory conditions (Ryan, 1992). She found that for nearly every experiment for which good data could be obtained, the fracture model employed in the SALE 2D code correctly predicted the fragment size distribution (Figure 4.1).

From this one can conclude that the fragmentation model has simulated the essential physics of the problem correctly. This allows one to apply this model to non-laboratory situations with some confidence in the result. This has already been done with respect to understanding the origin of the so-called Vestalets, 10 km-scale asteroids that spectroscopically appear identical to the asteroid 4 Vesta (Binzel and Xu, 1993; Asphaug, 1997). Application of an evolved version of this code showed how fragments of this size could be accelerated to Vesta escape velocity, resolving one of the final issues in explaining how the eucrites could travel from Vesta to the Earth. A key component of that work was developing a means to simulate fragmentation and ejection at high resolution, a theme to which I shall return. Sometime after the Vestalets paper was published, a crater of the size Asphaug predicted was discovered on Vesta by the HST (Asphaug, 1997).
4.2 Upgrades to SALE

I have added a number of improvements since I first worked with SALE. These upgrades add multiple material capability, greatly improved use of fictitious forces required for numerical stability, and corrects a long-standing logical error in that led to a phase mismatch between the strain calculation and the equation of state (cf. Melosh et al. 1992, Asphaug 1993).

4.2.1 Multiple material code

The SALE 2D code I use has been modified to incorporate multiple materials. A new array is defined to record the material number for each computational cell. Whenever the code needs to access a physical parameter for a cell, it reads the material number for that cell, then looks up the numerical value for a given parameter based on that material number. This approach allows the introduction of a single large array to track the material numbers for each cell and then very small arrays for each material parameter. This represents a great savings in computational time compared to the straight-forward approach of replacing each material parameter with an array as large as the entire computational grid.

This scheme was tested for self-consistency. For example, the grid was divided into rows or columns, each with a different material number. However those numbers all lead to identical material parameters. The results were exactly the same as the single material case, which is what one would expect. The first application of this capability was the study of atmospheric shock interactions with rocks on Venus (Head and Melosh, 1995).
4.2.2 Alternate node coupler

Numerical hydrocodes suffer from a problem known as "vertex coasting." During a calculation, it is possible for individual cells to change shape in such a way as to conserve volume. This appears as "bowtie" or "herringbone" patterns in the computational grid. Unchecked, these unphysical waves can lead to numerical instabilities as the distortion grows unabated. Eventually mesh entanglement occurs and non-physical volumes and energies result. Since the code has no way of "knowing" that materials resist shape change as well as volume change, a fictitious force is introduced to control or eliminate the phenomena. This is known as the alternate node coupler (ANC). The ANC examines the velocity of each node around a cell and adds a restoring velocity to counteract the short-wavelength distortions. The strength and mode of operation of the ANC is controlled by the user in the input file using parameters ANC and XI. Parameter ANC controls the strength of the fictitious restoring force while XI counteracts either bowtie or herringbone distortions depending on its value. Clearly it is best to use the smallest possible value for ANC since it is a fictitious force. A maximum value of ANC = 0.05 is recommended (Amsden et al., 1980) and was used in Ryan's and Asphaug's dissertation work (Ryan, 1992; Asphaug, 1993). In my work I use ANC = 0.001 and XI = 0.09. These values were chosen by trial and error in order to find the smallest possible of ANC that eliminated grid distortions related to vertex coasting.

Early in my work, I discovered a problem with the program logic implementing the node coupler. Since the fictitious velocity for any vertex depends on the velocity of each of the four vertices surrounding it, it is possible for the node coupler to generate a non-physical response of a vertex if the velocity gradient is very high. This is usually not a problem for intact homogeneous material. However, at the interface between two materials, especially if one of them is a fluid, this can lead directly to a non-physical solution.

In my simulations of atmospheric shock waves crushing rocks on Venus I
found that the rocks fragmented before the atmospheric shock actually reached it, and did so no matter how weak the applied shock pressure. The cause of this non-physical result I traced to the node coupler. I solved this problem by installing a logical switch in the appropriate subroutine. This switch turned the node coupler on or off based on the elastic properties and damage state of a given cell and its immediate neighbors. The altered code gave results that were much more reasonable—rocks did not break until the shockwave actually reached them and sufficiently weak shocks did not damage the rock at all. They merely blew them aside (Head and Melosh, 1995). Boris Ivanov has independently implemented an identical method for regulating the use of the ANC in SALE.

4.2.3 Strain calculation

While trouble-shooting some problems with the SALE hydrocode, I discovered a logical error in the use of the pressure calculation in the deviatoric stress calculator. Specifically, the problem arose in the calculation of tensional stresses that are used to determine if fracture occurs. There was a phase problem in that the pressure used in the stress calculation was from the previous timestep. This meant that for the first timestep, there was (erroneously) no confining pressure to counteract the deviatoric stresses, meaning that the first calculation of deviatoric stresses resulted in damage accumulation. By the second timestep the pressure was high enough to prevent fracture unless the stresses truly were tensional. However, for any calculation, the first timestep erroneously calculated a small amount of accumulated damage for cells next to the impactor. This is a non-physical result, purely an artifact of the phase mismatch between the pressure calculation and the strain calculation.

To repair the problem, I retarded the strain calculation. Instead of integrating the deviatoric stresses in the STRESD subroutine to get strain, I calculated the strain directly from the displacements in the REGRID subroutine. This left the computed pressure and strain in phase for the fracture criteria calculation, at
the expense of having no strain allowed during the first timestep. The new strain calculator was derived from first principles of the definition of strain and tested extensively against the original strain calculator, with good result.

4.3 High Resolution in the Spall Zone

Since the spall zone is roughly defined by the projectile diameter, the grid resolution in the zone can be characterized by the projectile resolution. In this regard the version of SALE described in Melosh et al. 1992 was inadequate. The projectile was created by taking several cells in the upper left corner and "pulling them out" of the mesh. These cells were given the desired impact velocity. This introduced a distortion that propagated into the mesh for a distance twice the projectile radius, i.e. a projectile three cells across (radius) distorted the grid 6 cells deep. Projectile sizes more than three cells across produce numerical instabilities that quickly render the solution untenable. This was acceptable in the earlier work that typically used projectiles a single cell across (Melosh et al., 1992; Ryan, 1992; Asphaug, 1993). However, this proved an unacceptable limitation for my study. If one considers an impactor 200 meters in diameter, then the maximum resolution achievable with this method is one-third of 100 meters or about 30 meters. Since the width of a shock in geologic materials is typically 1–10 meters, it was impossible to resolve the shock, let alone analyze its interactions with the surface.

A further limitation of hydrocodes appears at this point. Numerical hydrocodes do not reliably handle discontinuous features such as shockwaves. Strong shocks quickly lead to numerical instabilities. What is needed is a dissipative mechanism that spreads the shock over several cells (Anderson Jr., 1987). Viscosity is a dissipative mechanism. A method for efficiently introducing an artificial viscosity into the hydrodynamic equations (Von Neumann and Richtmyer, 1950) is now widely used. This means of course that the shock front is given an artificial thickness of about three cells. In the example above then, the shock front would be 90 meters
wide due to artificial viscosity alone, a thickness an order of magnitude larger than that due to material effects. Since spall velocity depends on the pressure gradient, any results using this resolution would be dominated by numerical artifacts, in this case the parameter governing the strength of artificial viscosity. Thus I needed a means of achieving much higher resolution in the projectile than had been attained before.

An additional problem presented itself as well. When the strength model was introduced to the original Los Alamos code (Melosh et al., 1992) a shear modulus was added to the calculation of deviatoric stresses. This provided the program with a means of "knowing" that cells were changing shape without changing volume. Since the ANC was designed to counter this characteristic, clearly there is potential for interference between these two features of the code. This was apparently not a problem for low-speed collisions (Asphaug, 1993). In my experience though, at high speeds the interference between the strength model and the node coupler results in unphysical results in highly sheared cells, i.e. those directly beneath the edge of the projectile. This problem is demonstrably exacerbated by the grid distortion the projectile simulation introduces. The problem manifested itself in the form of very large negative values for energy, and via the equation of state, pressure. This was limited to a few cells, but they were in the region where analytic calculations predicted the highest spall velocities. This was clearly an unacceptable situation.

4.3.1 Two step solution

To resolve these problems, I attempted to introduce master–slave surfaces to SALE, whereby the target and projectile grids are separate and interact across a defined boundary (Hallquist, 1978; Hallquist et al., 1985). During this effort, Boris Ivanov suggested a different approach that was much easier and proved satisfactory in nearly every way. He suggested a two step approach to the problem. First, the impact calculation is carried out in the Eulerian mode using a version of SALE
modified to simulate real, spherical impactors striking a target. Since this is an Eulerian calculation, material boundaries are poorly defined and fragment sizes cannot be calculated. Therefore, Lagrangian tracer particles are introduced into the calculation. These particles flow along with the material, recording the displacement and velocity histories for the selected points in the material. The tracer histories are recorded a file. This file is used to establish a boundary condition for a second calculation using SALE in the Lagrangian mode. In this way, the shock produced by the impact can be calculated in great detail in the Eulerian mode, then that same shock can be imposed on a Lagrangian calculation that has well-defined material boundaries and the capability to calculate fragment sizes.

I altered the boundary condition subroutines BC and BCSET in SALE to allow use of the tracer particle file. This new boundary condition examined the file at each timestep, "looked up" the velocity of each vertex at each timestep (there was a linear fit routine added since the timesteps cannot be expected to match) and impose this velocity on the left-hand boundary of the mesh. Typically the tracers defined a vertical line directly below the edge of the impactor. Thus the left side of the second calculation is offset one projectile radius from ground zero. This is why the simulation appears asymmetric. As a test, several runs were conducted in which the displacements of each vertex as a function of time in the Lagrangian calculation were compared with that in the Eulerian calculation. The displacements matched to within the limits of comparison.

As a result of this two-step approach to simulating the spall zone in an impact, I was able to increase greatly the resolution of the projectile. The previous limit was 3 cells across the radius. I have since carried out calculations with 40 cells across the projectile radius, an order of magnitude improvement. As I will discuss in Chapter 5, this resolution is sufficient to resolve the shock wave and examine its interaction with the free surface. This also means that my results are dominated by real physical effects, not by numerical artifacts.
The following sequence of Figures illustrate some of the output from a typical two-step calculation. The first step is shown in Figure 4.2. This is the Eulerian calculation. The Lagrangian tracers are shown, along with boxes delineating the boundaries of subsequent Lagrangian calculations. The velocity histories of the tracers define the velocity input for the second calculation. The results of that calculation are shown in time sequence for each of several output products: grid, velocity, pressure, energy, and damage. In each output product the propagation of the shock wave is evident, while various phenomena of interest are better discerned with different data sets. I have carried out approximately 75 computational runs of this type. This particular calculation produced no fragments that met the criteria for martian meteorites, as discussed in Chapter 3 and Chapter 5.

4.4 Conclusions

The SALE 2D hydrocode is well-suited for studying the general problem of spall in impacts and the particular problem of martian meteorite provenance. The code calculates fragment sizes with confidence. Modification allows the simulation of impacts into targets with realistic geometries at high spatial resolution. In doing so, boundaries between different materials and with the free surface remain well-defined. This code gave satisfactory results in a study of splotch formation on Venus. By utilizing the two stage approach, the modified SALE 2D hydrocode can be used to examine spall in impacts.
Figure 4.1: This is a summary of comparisons between the results from laboratory impact experiments and numerical models of each impact event. The agreement between the calculated and measured fragment size distributions is excellent (Melosh et al., 1992).
Figure 4.2: Step one of the two step method. The simulation is of a spherical projectile striking a target vertically at 10 km/sec. The dots in the target are Lagrangian tracers added to the Eulerian calculation. The box delineates the locus of grid points for the subsequent Lagrangian calculation. In that calculation the tracers define the left-hand velocity boundary condition for the computational grid. This approach was suggested by Boris Ivanov.
Figure 4.3: The computational grid used in the Lagrangian calculation. The left boundary condition lies directly beneath the edge of the projectile (compare Figure 4.2). The material parameters of the grid can be varied at will. For this simulation the grid is 50 cells across and 60 cells deep, i.e., 250 x 300 meters. The boundary condition was established by simulating a 100 meter diameter projectile impacting vertically at 10 km/sec. The crater is has a final diameter of 2.3 kilometers, deduced from $\pi$-scaling.
Figure 4.4: 11.5 msec into the simulation. The most distorted material is that closest to the projectile and displaced upwards. Note the absence of mesh entanglement.
Figure 4.5: 23 msec into the calculation. The progress of the shock wave can be discerned from the mesh distortion. Small instabilities are evident in the top row of vertices.
Figure 4.6: 34.5 msec into the calculation. Shock wave is approaching the boundary of the computational grid. The high-velocity spall has long since broken free. As a practical matter, all new spall calculations reported in this work cease approximately 30-40 msec after impact.
Figure 4.7: $T = 11.5$ msec. Plot of the velocity vectors. This Figure corresponds to Figure 4.4. Note that the vector length is rescaled in each of the following Figures. Separation of a spall layer has initiated. A single vertex has a speed in excess of Mars' escape velocity.
Figure 4.8: $T = 23$ msec. The velocity field reveals the propagation of the shock wave and the development of the excavation flow which will open the impact crater long after this simulation is complete. Note material separation near the surface.
Figure 4.9: $T = 34.5$ msec.
Figure 4.10: $T = 11.5$ msec. Pressure contours, units are pascals. "L" refers to low pressure. The "H" markings for high pressure have been deleted for clarity. The detached shock is readily discernible. Note that the contours "bend in" along the surface. No zero-pressure boundary condition is imposed upon the solution. The lower pressure near the surface is a consequence of shock wave interference.
Figure 4.11: $T = 23$ msec. Despite the "ringing" in the pressure contours, the propagation of the detached shock is apparent. Note that the lowest pressure isobar is closed off from the surface.
Figure 4.12: $T = 34.5$ msec. Note that the peak pressure has decayed by nearly a factor of two from the previous Figure. Negative pressure denotes material in tension that has not yet fractured.
Figure 4.13: $T = 11.5$ msec. Internal energy contours, units are J/kg. The energy contours closely parallel the pressure contours. Note that the minimum energy is slightly negative, but of a magnitude consistent with rounding errors.
Figure 4.14: $T = 34.5$ msec. Note that the internal energy near the surface is much lower than deeper material the same distance from the impact.
Figure 4.15: T = 11.5 msec. Damage contours. The material in the ejecta flap is completely fragmented.
Figure 4.16: T = 23 msec. Damage is propagating into the target material, but much more quickly along the surface. Analysis of the output data indicates that the maximum tensile stresses are sub-vertical.
Figure 4.17: $T = 34.5$ msec. Damage continues propagating into the target and along the surface. The instabilities apparent in the pressure and energy plots are evident here. An incipient secondary spall zone is apparent. Such secondary spalls were observed in the calculations by Ryan (1992) and Asphaug (1993). The near-surface material is fragmented upon ejection.
Figure 4.18: T = 35 msec. Isobars for the maximum pressure attained by a cell at any point during the computational run. The decay of shock wave is clearly evident, as is the near-surface interference zone. The highest maximum pressures are in the part of the mesh nearest the impact's equivalent center.
CHAPTER 5

Spallation

Spallation is a proposed mechanism by which impacts can accelerate intact target rocks to escape velocity. As noted in Chapter 2, it works because of interference between the free surface and the expanding impact shock wave, and is the means of meteorite ejection investigated in this work. This Chapter is devoted to the spallation process, how it is understood on theoretical and experimental grounds, and how it functions in different types of material with reference to the problem of martian meteorite ejection.

5.1 Analytic Spall Models

An analytical treatment of spall in impacts was developed by (Melosh, 1984; Melosh, 1985; Melosh, 1987). In the Melosh model, spall arises from the interference between the direct compressive wave and the tensile wave reflected from the surface (see Section 2.2.3). The compressive wave has a finite width, denoted rise time. Near the surface it is possible for a given location to sense the reflected tensile wave before the compressive wave has reached its full strength. The region where this occurs is denoted the interference zone. This zone is delimited by the locus of points where the tensile wave first arrives just as the compressive wave peaks. Below this line, the material experiences the full strength of the oncoming wave.

This line is defined by considering the arrival of the two waves from a common burst point at the equivalent center (one projectile diameter beneath the epicenter) of the impact. Equating the arrival times of the tensile pulse and the
peak of the compressive wave, the following equation can be derived:

$$z = \frac{c_l \tau}{2} \sqrt{1 + \frac{s^2}{d^2 - \left(\frac{a \tau}{2}\right)^2}}$$  \hspace{1cm} (5.1)

$Z$ is the depth to the interference zone as a function of epicentral distance $s$. The width of the wave front is $c_l \tau$, where $\tau$ is the rise time. Recall that $c_l$ is the longitudinal sound speed (Equation 4.3). The resulting curve is a concave-down hyperbola with a focus at the equivalent center $d$. This is the equation that Polansky and Ahrens (1990) fit to their experimental data.

### 5.2 Measuring the Interference Zone

According to the interference zone equation, the zone thickness is a minimum at $s=0$. Here, the interference zone thickness is $c_l \tau/2$. Recall the simulation of a stress wave in a one-dimensional rod. In that simulation the rise time was 10 msec and the sound speed was 3400 m/sec. The implied pulse width is about 34 m, meaning the interference zone should be 17 m deep at this point. It is evident in Figure 2.12 that the interference zone is half the width of the wave front. Examination of the output file shows that peak pressures attained during the run are reduced in the bottom 3 or 4 cells, i.e., the last 15-20 meters. I consider this to be a good match between theory and simulation.

One can apply this equation to the impact simulated in Chapter 4. In that case the projectile was 100 m in diameter striking at 10 km/sec. The width of the shock front in basalt ($c_l=6730$ m/sec) is 33.7 m. The interference zone varies from 17 m at the epicenter to 53 m at the right hand edge of the mesh 300 m from ground zero. The interference zone appears flatter than the low contour for maximum pressure in Figure 4.18. Examining the output file in detail indicates that the pressure gradient changes at the correct depth at 200 m and 250 m from the epicenter. The low pressure zone observed in the hydrocode output appears consistent with the Melosh model.
5.3 Parameter Studies

Those using hydrocodes are well-advised to go to great lengths to evaluate code behavior. An example of this was the computation I showed with the version of SALE used in this work to simulate 1D pulses in a thin rod. In that case the numerical and analytic computations agreed. Another example is the agreement between analytic and numerical calculations of the peak target pressure in the airblast calculation conducted in Head and Melosh (1995). In the case where no analytic solutions exist, one must rely on other means of code verification such as duplicating experimental results (Melosh et al., 1992; Ryan, 1992; Asphaug, 1993). When the subject studied is beyond the realm of experiment, then one should explore the parameter space to search for numerical artifacts. It has happened before that a reported science result from a numerical calculation was in reality a numerical artifact. Since first obtaining a scientifically interesting result, I have investigated code behavior to ascertain the degree to which those results can be considered reliable. In this section I discuss several of the studies I completed.

5.3.1 Resolution studies

One of the limitations of finite-difference hydrocodes such as SALE is the approximation inherent in rewriting the differential equations of motion in finite-difference form. This is characterized as the resolution of the simulation. There are at least two ways of defining resolution in hydrocode impact simulations. One criterion is the number of cells comprising the impactor (grid resolution). Another is the physical dimension chosen for each individual cell, given in this work in meters (cell resolution). These terms are related directly in that as one improves so does the other.

When I began working with SALE, the impactor (as described in Chapter 4) was defined by redefining the mesh boundary and cell sizes for a number of cells
in the upper left corner. Because of the distortions this introduced to the computational grid, the impactor size was practically limited to 3 cells across, meaning 6 cells in diameter once symmetry is taken into account. This caused great difficulty since it greatly limited the detail in which I could study spallation. In addition, adequate resolution studies were impossible to conduct since the impactor was either 1, 2, or 3 cells in radius. To evaluate the sensitivity of my results to grid resolution, I could only compare an impact event at three different cell sizes. For example, given a hundred-meter radius projectile, I could select cell sizes of 100, 50, and 33.3 meters.

This is bad for two reasons. First, for an impact of this size it would be impossible to resolve the shock wave separately from numerical effects such as artificial viscosity. Recall that in geologic materials the shock front is on the order of meters wide, while artificial viscosity spreads the shock over 3 cells, or at least 100 m in the example above. Secondly, any study of grid resolution effects is limited to a factor of three, smaller than the factor of 10 I examine as described below. The two-stage approach described earlier alleviated all of these difficulties. The chief limitation to resolution studies is now computer time.

Once I had used the two-step method to obtain good results (Head and Melosh, 1999b) I conducted a cell resolution study. I chose to simulate a 200 m basalt impactor striking a basalt half-space at 10 km/sec. These conditions were held constant while the cell size was varied from 20 m to 2.5 m. As a point of reference, the 1992 code was limited to a cell resolution of 33.3 meters for these impact conditions, i.e. the worst cell resolution used in this work was much better than the best achievable with the code as I received it. I recorded the maximum surface spall velocity as a function of distance from ground zero. The edge of the impactor then is at 100 m. This was done for cell resolutions equal to 20, 10, 7.69, 5, 4, 3, and 2.5 meters (see Figure 5.1). The results show a strong dependence of spall velocity on cell resolution for cell sizes comparable to or larger than the known thickness of the shock front in geologic materials. While for all simulations the spall
velocity declines with distance from the impactor, the velocities are never very high for low-resolution simulations (for cell sizes greater than 5 meters). Note also that the spall velocity declines monotonically with increasing cell size for simulation in this size range. This is exactly what one would expect for spall velocities determined by the pressure gradient when artificial forces define the shock width. In other words, in this resolution regime, the pressure gradient is determined by the cell size and the artificial viscosity, not by any real physical process. I conclude therefore that results from simulations with cell sizes greater than 5 m are incorrect, especially near the impactor, for reasons that appear well-understood.

Figure 5.1: Spall velocity as a function of distance. All physical parameters in the calculation are held constant. Only the cell resolution is allowed to vary. Note that spall velocity increases steadily as resolution improves until 5 cm/cell. Further increases in resolution do little to change the peak spall velocity. This is an indication that at 5 m/cell the grid is fine enough to resolve the impact shock wave.

This is in stark contrast to code behavior at relatively high resolution. Note that the results for cell sizes in the 2.5 to 5 m range overlap. To within a few percent, I get the same result even when the cell size is varied by a factor of
two. Furthermore, the monotonic change in spall velocity relative to cell size has disappeared. The only apparent size-related pattern is in the 2.5 m/cell case. There the spall velocities are systematically, though only slightly, less than the calculations for 3, 4, and 5 m/cell. Examining the distortion of the computational grid it is clear that the finer mesh suffers more from mesh entanglement—vertices nearer the impact pass through cell boundaries more removed. This means that energy is not being transferred as it should, resulting in higher calculated spall velocities near the edge of the impactor and lower values for vertices farther from ground zero. The remaining discrepancies do not have any obvious pattern. I interpret this to mean that these simulations are actually resolving the shock front and that numerical artifacts are not significantly influencing the result.

There remains another conclusion to be drawn from Figure 5.1. For the simulations at high resolution, the spall velocities are the same to within a few percent, but more importantly, they reach values in excess of Mars' escape velocity. In addition, much of this material is lightly shocked, to an extent comparable to that observed in the martian clan meteorites. The low resolution simulations (10 and 20 m/cell) do not show any material accelerated to escape velocity. The intermediate case shows a small amount of material accelerated to escape velocity, but greatly underestimates the peak velocities compared to the high resolution case. This observation provides one of the key conclusions of this entire work: in order to understand the ejection of intact material from Mars, one must examine the interaction of the shock wave with the free surface at very high resolution. This observation is generally applicable to spall studies on other planets.

In a further analysis, I have compared peak pressure as a function of distance from ground zero from a number of different computational runs (Figure 5.2). The input parameters are identical except for cell size. The peak pressure decreases with distance from ground zero and increases with depth as expected. Using data from three different calculations (cell sizes of 2.5, 4, and 5 meters) I can show that
the results from different runs are self consistent. The peak pressure increases mono-
tonically with depth when different runs using different cell sizes are intercompared. This further indicates that over this range of cell sizes the code is computing real pressures that depend on the physical system being simulated rather than fictitious forces.

Figure 5.2: Peak pressure is plotted against distance from ground zero at different depths in the computational grid. The horizontal line represents the shock pressures evident in the shergottites. The vertical dashed line denotes the outer limit of spall velocity exceeding escape velocity. Obtaining all the model data presented required repeated computations at different cell resolutions for the same impact conditions. The consistency between the different model results lends confidence that the hydrocode output reflects real physical effects and is not overly influenced by artifices.

Grid resolution is an important consideration in model data analysis. For example, I have simulated impacts of differing sizes into the same material using the same cell size. One impact generated spall in excess of escape velocity while the smaller impact did not. When I repeated the smaller impact calculation using a cell size designed to produce the same number of cells in the impactor as in the
larger impact case, the results were different. Several cells in the very top row exceeded escape velocity, meaning that both impacts liberated material from Mars. However, in the higher cell-resolution, small impact case, the liberated material all came from what would have been the top half of the cells in the lower cell-resolution case. Given this result, might there not be an even smaller impact that would launch material from the upper half of the high cell-resolution calculation? One can imagine that this is the beginning of an infinite, hopefully converging series. Fortunately geochemical data provides constraints that effectively rule out the ejection of material very close to the surface. Recall that no martian meteorite displays a two-stage cosmic ray exposure history. This means that none of the martian meteorites spent a significant period of time before liftoff within a meter or two of the surface. They must have come from somewhat deeper. Should a two-stage history be discovered in another martian sample, there is still the constraint of sample size. Clearly the spall layer must be thicker than the rock currently residing in a terrestrial museum.

5.3.2 Artificial viscosity

Artificial viscosity relates to the shockwave width as described previously. This has two effects on model data interpretation. The first is as described above, in that artificial viscosity can influence the pressure gradient if the cell resolution is too low. However, once high enough resolution is achieved, there is still the question of how the value chosen for the ARTVIS parameter may still influence the results. To investigate this, I conducted simulations that were identical expect that in the second step the artificial viscosity was varied by a factor of three. To evaluate the model data, I plotted pressure as a function of depth for several vertical slices into the computational grid at different distances from ground zero (Figure 5.3 through Figure 5.5). As can be seen, the results are much the same for very different values of ARTVIS. This is true even though ARTVIS was set to the unrecommended value of 0.4 (compare Figure 5.3 and Figure 5.4). When the cell resolution is varied by
a factor of two, the pressure gradients are indistinguishable (compare Figure 5.3 and Figure 5.5). This holds true at each distance from ground zero selected for analysis, which lends considerable weight to the interpretation of the cell resolution results above. If the artificial viscosity was dominating the shock wave structure, that should be evident in these Figures. Changing the value of ARTVIS should change the peak shock pressures in the interference zone. I found no evidence for such changes in my high resolution calculations. The amplitude of the instabilities (the ringing seen in the pressure plots) depends on cell size. For larger cells, they are smaller. As a matter of experience, these instabilities are largest at depth and practically disappear in the near-surface region.

![Peak Pressure at Depth](image)

Figure 5.3: Plot of pressure against depth for different vertical slices into the target. 200 m diameter impactor at 10 km/sec with 2.5 m/cell resolution. ARTVIS = 0.2. The instabilities are largely a product of the small cell size.

The second difficulty with artificial viscosity is that even though the shock front is being resolved properly, the actual pressure in any individual cell in the interference zone may still be effected. That there may be unexplained cell-to-cell variations in pressure is actually evident in the artificial viscosity Figures in the form
Figure 5.4: Same impact conditions as the previous Figure, but with ARTVIS set to 0.4. There is little change in the pressure structure.

of numerical instabilities. As a rule, the artificial viscosity routine smears the shock over three cells. Hence, one should be cautious when interpreting output data of peak shock pressures near the surface. Theoretically, the pressure at the free surface is always zero. However, pressure is a cell–centered quantity and therefore pressure at the surface itself cannot be calculated. The best one can do is calculate the pressure at a depth one–half the cell resolution. In my simulations, the shallowest pressures I can calculate are for a depth of 1.25 m. This individual cell value for pressure is certainly influenced by artificial viscosity at low resolution. This effect has been neglected in the literature. A particularly bad example was an attempt to calculate surface pressures antipodal to large impacts on Mars (Williams and Greeley, 1991). In that calculation the cell size was greater than 100 km.

To estimate the peak pressures encountered by impact ejecta, I considered the following. The pressure at the surface is certainly less than the pressure 3 cells down. The pressure probably does not decrease linearly as one approaches
Figure 5.5: Same conditions as Figure 5.3 but with resolution set to 5 m/cell. The primary difference is the sharp reduction in numerical instabilities.

the surface, but more likely declines slowly until one is quite near the surface, then decreases rapidly (Melosh, 1984; Melosh, 1987). In addition, recall from the discussion of CRE data that martian meteorites contain no evidence for a two-stage exposure history. This constrains the depth of burial on Mars to greater than a meter or so, more likely 2–3 meters. Hence, model pressures nearer the surface than 2–3 meters are probably irrelevant in examining this problem.

5.3.3 Weibull parameters

The fracture model used in SALE begins with a model for the distribution of flaws in a material. The flaw distribution is governed by the so-called Weibull parameters $k$ and $m$ as discussed previously. Numerical values for these parameters in geologic and construction materials are summarized in Table 5.1. To test the robustness of my results with respect to the Weibull parameters, I conducted a series of simulations.
where \( k \) was varied by several orders of magnitude from \( 10^{29} \) to \( 10^{37} \) and \( m \) was varied by up to 50%. The principle results were not affected by varying these parameters over this range of values. The fragment sizes and peak spall velocities were essentially unchanged. Hence, I consider my results to be robust with respect to these parameters. Other factors have a greater influence over fragment size and spall velocity.

<table>
<thead>
<tr>
<th>Material</th>
<th>( k ) ( (m^{-3}) )</th>
<th>( m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Westerly Granite</td>
<td>( 4.15 \times 10^{23} )</td>
<td>6.2</td>
</tr>
<tr>
<td>Basalt</td>
<td>( 1.0 \times 10^{33} )</td>
<td>9.5</td>
</tr>
<tr>
<td>Alluvium</td>
<td>( 4.15 \times 10^{23} )</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Table 5.1: Weibull parameters for materials simulated in this work. The parameters for alluvium are for convenience. I do not use competent alluvium as a proxy for any martian terrain. Rather, I use damaged alluvium. In that case the material is already fragmented and the value of the Weibull parameters is moot. Data are from Ryan (1992), Melosh et al. (1992), and Asphaug (1993).

5.3.4 Material properties

Most of the material properties used in the hydrocode simulations are manifested by the Tillotson equation of state (Tillotson, 1962). The parameters used in this work are summarized in Table 5.2. Equations of state (EOS) relate pressure, density, and internal energy. Everyone is familiar with the Ideal Gas Law \( PV = nRT \), the simplest EOS requiring a single measured parameter \( R \). Obviously it is useful only in the approximation of systems of interest. In detail it is inadequate. For example, an Ideal Gas cannot be refrigerated, nor can the somewhat more complicated van der Waals gas (Reif, 1965).

State-of-the-art EOS can be quite cumbersome. The ANEOS equation of state package is not really an analytic EOS but rather a lookup table—a FORTRAN code comparable in length to SALE2D. It has capabilities well beyond the Tillotson, at the price of great complexity (for comparison purposes, the IGL and the VDW EOS use 1 and 2 parameters respectively. ANEOS uses about 40). Fortunately
Table 5.2: The Tillotson parameters used in this study for Westerly Granite, basalt, and Nevada Test Site alluvium. Data are from compilation by Melosh (1989). The basalt data is really "anorthositic gabbro," but is identical to the parameters used in Ryan (1992).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Granite</th>
<th>Basalt</th>
<th>Alluvium</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho (\text{kg/m}^3)$</td>
<td>2680</td>
<td>2820</td>
<td>2600</td>
</tr>
<tr>
<td>$a$</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$b$</td>
<td>1.3</td>
<td>1.5</td>
<td>0.8</td>
</tr>
<tr>
<td>$A$ (GPa)</td>
<td>18</td>
<td>71</td>
<td>30</td>
</tr>
<tr>
<td>$B$ (GPa)</td>
<td>18</td>
<td>75</td>
<td>10</td>
</tr>
<tr>
<td>$E_0$ (MJ/kg)</td>
<td>16</td>
<td>487</td>
<td>6</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$\beta$</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$E_{iv}$ (MJ/kg)</td>
<td>3.5</td>
<td>4.7</td>
<td>3.5</td>
</tr>
<tr>
<td>$E_{ic}$ (MJ/kg)</td>
<td>18</td>
<td>18.2</td>
<td>18</td>
</tr>
</tbody>
</table>

A detailed thermodynamic history of each cell ranging from solid rock to vapor is not of concern in this work. A typical run time is about 30 msec, by which time material near the impact site has fragmented and been launched from the surface. The Tillotson EOS is a good compromise and has been used with great success in the SALE2D code (Ryan, 1992; Melosh et al., 1992; Asphaug, 1993).

I varied several of the material properties again to test the robustness of my results. By far the most important of these was damage. Damage $D$ is calculated cell by cell throughout the computation run. Intact material has $D = 0$ and fully damaged material (strengthless in tension) has $D = 1$. It is possible to assign an initial value for damage on a cell by cell basis in part because of the hydrocode's added multiple material capability. This facilitates examining the effects of layering discussed in the next section. By assigning $D = 1$ to all the cells in the target I am effectively reducing the shear modulus to zero in tension. This reduces the longitudinal sound speed $c_l = \sqrt{(K + (4/3)\mu)/\rho}$, where $K$ is the bulk modulus, $\mu$ is the shear modulus, and $\rho$ is the density. Obviously, reducing $\mu$ reduces $c_l$. In simulation, it is evident that the shock wave progresses less rapidly through the mesh when damage is preset to zero as expected. In addition, the spall velocity
is much reduced as well, typically about 20%. This is a general result, present for every material substance I tested. The drawback of these calculations is that fragment sizes in initially damaged cells are undefined.

I am interested in the effect of initial damage on the propagation and strength of the shock because damage is a means of simulating a martian regolith. One can imagine that a fresh surface on Mars may be composed of a lava flow. Over time, weathering processes would tend to weaken the material through cracking, frost wedging and the like. Eventually an initially intact lava flow would be reduced to an organized collection of boulders. In addition, impacts over the eons should fragment and even pulverize the surface. Certainly there is abundant evidence for fine-grained mobile material on the martian surface as well as boulders and ejecta blocks. Given these processes, the thickness of the regolith (the depth to which the surface material is fractured and mixed) may well be proportional to age. Hence, initial damage may be an important parameter in characterizing the differing response of martian surface materials to impacts.

Another important parameter is crack speed. Crack speed is generally considered to be about 40% of the longitudinal sound velocity (Irwin, 1958; Jaeger and Cook, 1969). The crack speed strongly influences the fragment size calculation, in the sense that the higher the crack speed the larger the fragments. This may appear counterintuitive, but it is related to crack coalescence. When flaws are activated they are assumed to grow at the crack speed until the flaw is no longer active (typically at coalescence). Since the growth of cracks shelters stronger flaws from the strengthening stress field, a higher crack speed means that a greater volume of stronger flaws is sheltered than in the case of a lower crack speed, meaning that fewer flaws will be activated. Since there are fewer active flaws, the material is broken into a smaller number of larger fragments. The relationship between crack speed and fragment size is linear, so that doubling the crack speed doubles the fragment size, all other variables held constant (Melosh et al., 1992).
5.4 Layering Effects

Layering is an important consideration for understanding the effects of impacts into realistic geologic targets. In our explorations of the solar system it appears that nearly every planetary body studied is layered to some extent. In every case for which there is sufficient data the planet is covered by a regolith. Certainly this is true for Mars. Our understanding of the extent of layering on Mars is undergoing something of a revolution now as the highly detailed images from the Mars Global Surveyor become available. Recall that on Mars there appears to be no mechanism at work recycling crustal material. Once rocks form on the surface, they appear destined to remain there, possibly to be covered by a later lava flow or ejecta blanket, but never to be subducted back into the mantle. Erosional processes are probably orders of magnitude weaker than on Earth. Thus, thick sequences of layers on Mars should be the expectation.

I have simulated impacts into layers composed of two different materials: basalt and alluvium. Most of the basalt parameters are derived from Ahrens and O'Keefe's work to develop EOS parameters for a fictitious rock type termed "gabbroic anorthosite." The Weibull and Grady-Kipp parameters are taken from experimentally derived values for basalt. I use basalt because that is the composition for most of the shergottites— they are rocks formed in basaltic lava flows on or very near the surface of Mars.

I use the EOS parameters for alluvium as an analog for the martian regolith. This is a well-studied material, the data collected from years of nuclear and high-explosive testing at the Nuclear Test Site in Nevada. There are fundamental differences between a regolith and lake alluvium. The martian regolith almost certainly has not experienced the types of lithification processes common on Earth. It is doubtful that abundant liquid water flows through the material on a regular basis, dissolving and cementing the material into a coherent rock. There is some
evidence from the three martian lander missions that the regolith has an undulating character, perhaps indicative of impact-related diagenesis. This is admittedly speculative. However, until we can obtain good data about the martian regolith (local seismic studies or drill cores) reasoned speculation is all that is available to constrain the physical character of the surface material. To simulate the regolith, I typically use the EOS parameters for alluvium with the damage preset to 1. That is, I consider the martian regolith to be a strengthless version of terrestrial alluvium. When better-motivated data becomes available, I shall certainly use it. For the time being, this approximation must be considered during model data analysis and interpretation.

To simulate impacts into layered terrains, I use the gabbroic anorthosite as an analog for martian crustal material (bedrock) overlain by damaged alluvium (regolith). I have conducted a systematic series of calculations to examine the effect of the layered structure on the shock wave. I have considered the effect of layer thickness and examined this with respect to spall velocities. In the most complicated case to date I have used a three layer model: intact basalt (a recent lava flow) over damaged alluvium (regolith) over intact basalt (bedrock).

5.4.1 Alteration of the shock wave

The seismic velocity for alluvium, especially damaged alluvium, is considerably less than that of basalt. As one might expect, the shock wave is retarded in the low-velocity layer relative to the higher-velocity bedrock. This results in a distortion of the shock wave as can be seen by comparing Figures 5.6 and 5.7. Besides the obvious distortion of the shock wave with respect to the case of an impact into an infinite half-space, there are two important effects to note. First, the shock morphology near the surface has been altered in such a way as to make the near-surface pressure contours more parallel with the surface. Second, the strength of the pressure gradient is significantly enhanced in the layered case—about 50% in
the figure presented. This has the overall effect of altering the spall velocity of the near-surface materials, increasing it over the values obtained for impacts into an alluvium halfspace.

Figure 5.6: The figure shows a shock wave in a homogeneous half space composed of basalt. The shock wave is hemispherical with respect to the equivalent center 200 meters deep and 100 meters left of the mesh boundary.

5.4.2 Spall velocity vs. distance

As one would expect, the peak spall velocity decreases with distance from the edge of the projectile. This phenomenon was analyzed quantitatively by Melosh (1984). In that work, Melosh derived a velocity falloff rate of $1/r^2$. In my hydrocode simulations, I have not witnessed velocity falloffs as calculated in earlier work. In
Figure 5.7: Identical to Figure 5.6 except the upper 50 meters is composed of a low-velocity material (damaged alluvium). Near the surface the shock wave lags behind the deeper shock. Note also the differences between the two cases of the near-surface shock morphology and strength of the pressure gradient.

Simulations of impacts into infinite half-spaces I have observed velocity falloffs of about 2 in the exponent (Figure 5.8). This is a somewhat less steep decline. The possible reason for the difference comes from an incorrect theoretical model for the width of the shock front. In Melosh (1984) the rise time of the shock wave was equated to the target penetration time of the projectile $t = v/d$ where $t$ is the rise time, $v$ is the impact velocity, and $d$ is the projectile diameter. As discussed previously, this formulation is used to derive the locus of the interference zone. A prediction from this model is that point source explosions should have no interference zone (see Section 5.1). This is not the case as observed in the nuclear test
data (Section 2.2.3). Melosh has argued that the width of the shock front may be a material property, rather than tied directly to impactor diameter (Melosh, personal communication). This could be the reason for the difference.

For impacts into layered terrains, the spall velocity falloff can be quite different from the infinite half-space simulation. In general I have simulated impacts into layered terrains where a low-velocity layer overlies a higher velocity layer (regolith over bedrock). Sometimes the regolith layer is topped by a thin high-velocity layer (the recent lava flow). In all of these cases the spall velocity decreases with distance from ground zero even more slowly than before (Figure 5.8).

![Figure 5.8: Spall velocities as a function of distance from ground zero, for different target compositions. Note that in the layered terrain calculations, the velocity fall off is less steep than in the homogeneous case.](image)

The observation that layering can influence not only the value of the peak velocity but also its radial variance has possibly important implications for the selective liberation of material from the parent body, Mars. It is clear from a perusal of Figure 5.8 that there exists a regime where identical impacts into layered
vs. homogenous terrain would eject a much greater volume of material to escape velocity. From the Figure it appears a planetary escape velocity of 3.5 km/sec would be in the critical range for the impact simulated. One can imagine a higher-velocity impact where the critical velocity would be approximately Mars' escape velocity of 5 km/sec. The implication of this is that impacts into layered terrains should be much more efficient in ejecting material from Mars. Such terrains might be typified by recent lava flows over older regolith. Thus, material from recent lava flows might be overrepresented in the collection of martian meteorites for this reason alone. To date I have not been able to simulate impacts large enough to observe this effect in detail for critical speeds near Mars' escape velocity. This inability is primarily a limitation of currently available computer speed. Note that the critical velocity in Figure 5.8 is approximately lunar escape velocity.

The effectiveness of the regolith in reducing spall velocity decreases steadily with increasing depth until the regolith layer is approximately as thick as the impactor diameter (Figure 5.9). The effect of increasing regolith thickness is to reduce the spall velocity nearest the impact site first. Since this is the location of ejected material, a much thinner layer, on the order of 1/6 the projectile diameter, can suppress launch of meteorites.

The damage state of target material has a strong influence on peak spall velocities. Simulation of impacts into damaged material was accomplished by the simple expedient of setting the damage parameter to 1 in the CELSET subroutine. In the stress deviator subroutine, material that has been damaged has the shear modulus reduced to zero, which causes a corresponding reduction in the longitudinal sound speed. For all simulations of impacts into damaged materials, the spall velocities were reduced compared to impacts into competent material (Figure 5.10). The reduction in spall velocity is evident at all distances from ground zero. This result is consistent with scaling laws for crater diameter. For identical impact conditions, the final crater size is smaller for impacts into sand as opposed to impacts into competent rock. Impact energy in former case is diverted from the excavation
5.5 Spall Thresholds

Of great interest to me is determining as precisely as possible the size of crater required to eject the martian meteorites. Once this size is well-characterized, then I can place the result into the larger context of the various geochemical, remote sensing, and dynamical information presented heretofore. This information has provided a number of constraints on plausible martian source craters. First of all, material must be ejected to escape velocity, which is 5.0 km/sec on Mars. A related note is that the required velocity for direct injection into an Earth-crossing orbit is 5.35-5.9 km/sec, depending on whether Mars is at aphelion or perihelion respectively at the time of impact. It is not enough that the vertices along the surface record
Figure 5.10: The figure shows spall velocity as a function of distance from ground zero for two impact simulations. The impact conditions are identical except for the initial value of the damage parameter. The damaged material (D=1) has a systematically lower spall velocity than the undamaged material (D=0). This result holds for all materials simulated.

speeds in excess of escape velocity. Escape velocity must be achieved at some finite depth. The criterion I use is that for any cell, both lower vertices must have spall velocities in excess of the escape velocity to consider the material in that cell to be ejected.

The material must be in the size range of the martian meteorites, with proper allowances for atmospheric ablation. The fragments cannot be so large as to give rise to a large volumetric fraction of material with $2\pi$ CRE. The material must come from a depth of two or more meters, again because of the complete lack of evidence for $2\pi$ CRE. The peak shock pressures for the shergottites are generally 30-50 GPa, for Chassigny about 35 GPa, and for the nakhlites undetectable, meaning less than 5 GPa based on Table 2.1. Hence, peak shock pressures recorded for the ejected fragments must fall within that range. The impact should liberate a
sufficiently large number of fragments ($\sim 10^6$) that one would expect at least one to reside in our martian meteorite collection. Ideally, I would want $\sim 10^7$ fragments to allow for the observed pairing in the lherzolites and basaltic shergottites.

Finally, the minimum crater should be established for different target compositions and stratigraphy—of especial importance is the damage state of the material prior to impact. The simulated target should conform to best understanding of martian geology. To this end, I use a basalt analog for the martian surface and damaged alluvium for the regolith. Since the chief means of regolith production on Mars is probably impact cratering, the depth of the regolith is probably correlated with crater age. The depth can be estimated from the depth of craters for which the surface is saturated. For young terrains, only the smallest craters are saturated, so the regolith is thin. For older terrains, the saturation crater size is larger and the regolith is correspondingly deeper. In addition, I consider the possibility that recent lava flows could overlie weathered terrain. This results in a three-layer model.

5.5.1 Competent material

Impacts into competent material means that the surface layer is undamaged prior to impact. I include in this category impacts into layered terrain where one or more of the layers is pre-damaged. A number of impact scenarios produced candidate shergottite material. The smallest event simulated the impact of a 150 m diameter projectile at 10 km/sec. The $\pi$-scaled final diameter for this event is 3.1 km. I chose target materials of a basalt halfspace and layered terrain composed of basalt over damaged regolith. The greatest volume of material ejected to escape velocity was for the basalt halfspace target. I analyzed output data for three different values of cell size for these impact conditions. For the 3 m/cell case, material was spalled to escape velocity to a depth of 3-6 meters in an annulus extending 153 to 180 meters from ground zero. Examining the output data on a cell by cell basis, I determined that 2 cells in the lower layer and 7 cells in the topmost layer contained fragments shocked
to shergottite-like pressures with sizes comparable to some shergottite samples. The impact ejects a total of $7 \times 10^7$ rocks, the vast majority of which (99%) are 3-10 cm in diameter. In the analysis I excluded material from the top two meters of the target to meet the CRE constraint.

A similar analysis of the same impact with a cell size of 3.75 meters gives a similar result. Approximately $4 \times 10^7$ fragments are ejected with a size range of 5-9 cm. One reason for the lower yield is that only a single layer of cells exceeds escape velocity as opposed to two layers in the previous case. These results are consistent in that they indicate the maximum depth of ejected spall is between 6 and 7.5 meters.

The picture is complicated by the analysis of the same impact with a resolution of 5 m/cell. The inconsistency here is the fact that no vertices other than those on the surface exceed escape velocity, indicating a maximum depth of spall of less than 5 meters. This lower depth may be a resolution effect. Recall that there was very little difference in spall velocities for cell resolutions ranging from 2.5 to 5 m/cell. The slight discrepancies apparent in that study are perhaps evident here in the detailed analysis. By interpolating the spall velocities with depth on a vertex-by-vertex basis I can still estimate a volume of ejected material. The yield in this case is approximately $3 \times 10^7$ fragments in the 4 cm size range. All these fragments come from a single cell with a peak shock pressure somewhat too high for shergottites ($\sim 60$ GPa), so this lower resolution analysis serves mainly as a check on the previous analyses. All three indicate the same order of magnitude in number of ejected fragments.

The number of fragments from this event is sufficient to expect several to reside in our collections. The fragment sizes are in the 5-10 cm range, which is consistent with the sizes of the lherzolitic shergottites ALH77005, LEW88516, and Y-793605. Recall that these three are source crater paired from geochemical analysis (Warren and Kallemeyn, 1997).
<table>
<thead>
<tr>
<th>Fragment Size (cm)</th>
<th>Number</th>
<th>Shock Pressure (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2 \times 10^9$</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>$6 \times 10^7$</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>$4 \times 10^8$</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>$8 \times 10^6$</td>
<td>65</td>
</tr>
<tr>
<td>15</td>
<td>$5 \times 10^5$</td>
<td>35</td>
</tr>
<tr>
<td>26</td>
<td>$4 \times 10^5$</td>
<td>$\leq 30$</td>
</tr>
</tbody>
</table>

Table 5.3: Analysis of 150 m diameter, 10 km/sec impact into a homogenous basaltic terrain with a resolution of 3.0 m/cell. The majority of fragments are from a single cell with a calculated mean fragment size of 1 cm. These fragments are too small to be viable martian meteorite progenitors (see Table 3.1. The largest fragments are comparatively rare. It is unlikely the Zagami-sized fragments from this impact are numerous enough to expect to find them on Earth (see Section 3.5.1).

I analyzed results from a simulation of a 200 meter impactor striking a basalt target at 10 km/sec at 5, 3, and 2.5 m/cell resolution. The $\pi$-scaled size of the final crater is 3.9 km. The results were internally consistent in that the number of ejected fragments in each case was a few times $10^7$ from the same size spall zone. The width of the ejected spall at a depth of 5 meters was 30 m and 27.5 m for the 5 m/cell and 2.5 m/cell runs respectively. The width of the ejected spall zone at 6 m depth was 21 m. The width of the ejected spall at the surface was the same within a single resolution element. This consistency in the detailed analysis lends confidence to the results. All three analyses indicate that about $10^7$ fragments in the 5-15 cm size range are ejected. Approximately $10^6$ fragments in the 15-20 cm size range and about $10^6$ fragments in the 25-50 cm size range are ejected. The number of fragments and their sizes are consistent with the basaltic shergottites (Shergotty, Zagami, and QUE94201).

Comparing the 5 m/cell and 2.5 m/cell runs shows that the peak shock pressure in the uppermost row was uniformly less by a factor of two in the higher resolution run. Peak shock pressures in the second row of the high resolution run were uniformly about 10% higher than the top row of the 5 m/cell run. Again, the peak pressures are internally consistent for a given impact event studied at different resolutions. This suggests that the undetectably low shock pressures inferred for the
nakhlites might be reproduced at still higher cell resolution, though this approach may have difficulty meeting the CRE data.

<table>
<thead>
<tr>
<th>Fragment Size (cm)</th>
<th>Number</th>
<th>Shock Pressure (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-10</td>
<td>$1 \times 10^7$</td>
<td>40-50</td>
</tr>
<tr>
<td>12-15</td>
<td>$1 \times 10^7$</td>
<td>20-65</td>
</tr>
<tr>
<td>17-21</td>
<td>$2 \times 10^6$</td>
<td>20-60</td>
</tr>
<tr>
<td>27-28</td>
<td>$2 \times 10^8$</td>
<td>30-50</td>
</tr>
<tr>
<td>37</td>
<td>$1 \times 10^5$</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 5.4: Analysis of 200 m diameter, 10 km/sec impact into homogenous basaltic terrain with a resolution of 2.5 m/cell. In this run, the minimum mean fragment size for any cell was 5 cm. This is the main reason the number of fragments ejected is a factor of ~6 smaller than the smaller impact described above.

For the same impact into a layered target, the result is different. The target is composed of a single 2.5 m thick top layer of basalt overlying 40 m of damaged alluvium overlying a basalt halfspace. The width of the ejected spall zone at the surface roughly the same as before, 45 m. However, the spall velocities at depth decline very quickly. There are no vertices in excess of escape velocity 5 m below the surface. For the homogeneous target, the ejected spall zone was 30 m across 5 m deep. The yield from this simulation is $3 \times 10^6$ fragments in the 7-10 cm range. The same calculation at 5 m/cell resolution revealed no escaped fragments, a consistent result.

Martian craters of diameter 3-4 km represent the lower bound for plausible martian meteorite source craters. These simulations show that impacts producing such craters also eject to escape velocity fragments of sufficient size, shock state, and in large enough numbers to account for the shergottite samples in our collections. The CRE data indicated 4 impacts for the shergottites with the samples arranged into two groups of three and two singles. The events producing a single sample on Earth are not necessarily smaller than the other two because they may be more recent. The CRE age for DAG 476/489 is one-third to one-quarter that of the lherzolites or the paired basaltic shergottites. The CRE age for EET79001 is three times smaller still. It takes approximately 10 Ma years to deliver half the samples
that eventually find their way to Earth, hence the impacts responsible for EET79001 and DAR476/489 most likely have other representatives that will be collected in the far future. For this reason, I do not ascribe a smaller impact as the source of these samples.

Recall the demonstration in Chapter 4. That impact event resulted in a \( \pi \)-scaled diameter of 2.3 km. It is not a good source crater candidate for the known martian samples, despite the fact that it did liberate \( 3 \times 10^7 \) fragments in the 4 cm size range. Three of the martian meteorites had a pre-atmospheric size of about 3 cm, however each of these is source crater paired with larger fragments (see Table 3.1). The Iherzolites are the closest match, with ALH77005 being about 8 cm across in space. The implication here is that smaller impacts can eject material from Mars, even if we do not have any samples of such an event. It is clear that the 3.1 km crater was large enough to be the source for the Iherzolites, while the 2.3 km event is probably too small. Therefore I adopt 3 km as the minimum crater required to eject the known meteorite samples from Mars. Naturally this limit can be revised downwards in the event small, unpairable samples are discovered.

5.5.2 Damaged material

Impacts into damaged material means that the target, from the surface to a depth comparable to the impactor radius, is fragmented before the simulation begins. In general, spall velocities are reduced compared to similar impacts into competent rocks (Figure 5.10). This means that larger craters are required to accelerate material to escape velocity. I investigated the impact of a 200 m diameter object at 10 km/sec into a terrain composed of regolith of various thicknesses overlying basaltic bedrock. I cannot calculate a fragment size in previously damaged material—I can only determine if that material would be ejected to escape velocity. For comparison purposes, this impact into homogeneous basalt (at the same cell resolution) produced about \( 5 \times 10^7 \) fragments in the 10 cm size range. For a regolith depth of 200
m, only a single cell exceeds escape velocity. However, the peak shock pressure in that cell is calculated as 70 GPa. In general, shock pressures near the surface are much higher in simulated regolith than in competent material. This is reasonable because the calculated particle velocities are lower. Conservation of energy requires some sort of compensation for these lowered velocities, hence, higher shock pressures and internal temperatures. A similar result holds when the regolith layer is reduced to 30 meters. In that case, two cells exceed escape velocity, but with shock pressures of 70 and 84 GPa. This material is too highly shocked to be martian meteorite material. Thus 4 km craters into terrain covered by just 30 m of regolith are not viable candidate source craters. Since the number of impactors varies inversely with size, the efficiency of spall ejection decreases with regolith depth, i.e., terrain age.

Doubling the impactor size does generate possible martian meteorites. For a 400 m diameter impactor at 10 km/sec into a 100 m layer of regolith over basalt, approximately $10^6 m^3$ of material is ejected with shock pressures ranging from 40-60 GPa. This is at the upper limit for shergottite shock pressures. Assuming that all the fragments are 30 cm across, the impact ejects $4 \times 10^6$ fragments into space. This is a sufficient number to expect some samples in our terrestrial collection. The resulting crater is 6.7 km across. Given that the shock pressure constraint is barely met, I consider this to be the lower limit for impacts into older terrains. This is viable for shergottite material, but not nakhlite or Chassigny. For these, larger impacts would be required to meet the shock pressure constraint.

A regolith layer overlying bedrock reduces the spall velocities. The reduction increases until the layer is approximately 1/5 to 1/4 the diameter of the projectile over the size range I have investigated (impactors 100 m to 400 m in diameter). This obviously suggests a trend that might continue to ever increasing regolith depths, i.e., terrain ages. The implications this has will be addressed more fully in Chapter 6.
5.5.3 Speculations on porosity

Porosity is a material property I do not model in SALE, other than peripherally in the EOS parameters. For example, alluvium is a porous material, and the EOS parameters determined in the laboratory must reflect that to some extent. However, an explicit treatment of porosity resists a satisfactory solution at present. Research groups at the national laboratories (Las Alamos, Livermore) and in academia (Asphaug, Ryan, Melosh) have not yet constructed a model of porosity that works well. Ideally, we would develop a porosity subroutine that could be dropped into the various programs in use. More likely, a good model for porosity will require a re-evaluation of the distribution of energy in impacts. This could mean a substantial change in the energy subroutines used in hydrocodes.

Porosity is important because it is a good way to dissipate energy and we have very good reason to believe that many targets of interest are very porous indeed. The images from the flyby of asteroid 253 Mathilde show 5 giant (larger than the asteroid's radius) craters (Veverka et al., 1999). These craters appear to abut one another, as if they formed without disturbing their neighbors. Apparently, the expanding shock wave dissipated very quickly, resulting in negligible particle velocities outside the crater. The same effect has been demonstrated in laboratory impacts into highly porous targets (Housen and Holsapple, 1999). This should motivate future experiments aiding the development of models for porosity which can then be applied to interpreting observations of astronomical objects.

Until porosity can be modeled in a reliable fashion, one can only speculate as to its possible effects on spall efficiency. It seems likely that porosity will continue the trend observed in my work. Impacts into competent material produces the highest spall velocities. Impacts into competent but less dense material has reduced spall velocities, while impacts into damaged materials have the lowest spall velocities modeled. Given the addition sink for impact energy in the form of PdV work (crushing pore spaces) I would expect spall velocities to be even lower. Certainly
the excavation flow appears highly attenuated in laboratory impacts into porous materials (Housen and Holsapple, 1999) and in apparently porous asteroids. Hence, the limits established above for the minimum required crater to eject material from martian regolith are probably strong lower limits. The actual required sizes may well be larger.

5.6 Conclusion

I have employed a new method for analyzing spall in impacts. By making the calculation twice in succession, I am able to simulate the impact with high resolution in the interference zone while retaining the capability to estimate fragment sizes and shock state. This allows one to calculate the number of fragments with velocity in excess of escape velocity for Mars.

The relevant parameters in the calculation have been examined to determine their effect on the results. Artificial forces do not appear to play any significant role. Thus I have been able to examine the effects of variable target compositions and physical state on spall efficiency. This allows me to determine threshold crater sizes that are target-dependent as well as impactor energy and velocity dependent. Before this work, the question regarding ejection of meteorites from Mars centered on the minimum size impact required. The high resolution study gives a result much smaller than previously thought, at least for competent targets. Additionally, the idea that the minimum size impact depends on target composition and physical state adds a new dimension to the analysis. Consequently, I am now able to postulate scenarios for martian meteorite provenance more sophisticated than any previously possible.
CHAPTER 6

Scenarios for Martian Meteorite Launch and Delivery

In this Chapter I present scenarios for the fragmentation and ejection of martian meteorites. These scenarios take into account all the constraints discussed in the previous Chapters, including results from geochemistry, orbital dynamics, and my own hydrocode simulations. It is my intention that this Chapter, so far as it is possible to do so, posits the definitive version of events leading to the arrival of martian meteorites on earth.

6.1 Impacts into Fresh Material

Fresh material is young, meaning a minimum of impact gardening has fragmented the rock. Impacts into this terrain would produce the shergottites. My conceptual model is of a young lava flow, presumably competent basalt of some thickness. It is possible that the flow or sequence of flows is on the order of 10s of meters thick, overlying older terrain that might be weathered, perhaps heavily weathered. Impacts in the 7-12 km/sec range producing craters 3-4 km in diameter or larger are viable source craters, provided the impacts are into targets of this type. Such craters are abundant on Mars.

The lherzolite-liberating event requires a crater only 3 km across (Section 5.5.1). Because this event generates many fragments much larger than observed in the lherzolites, it is possible that the actual source crater was smaller. This conclusion may be adjusted if an additional lherzolite sample larger than ALH77005 and paired with it is found (Table 3.1). The basaltic shergottite-producing craters had
to be somewhat larger than the minimum postulated lherzolite-producing crater because of the larger required fragment sizes. Craters 4 km across meet this constraint. Before detailing a chronology for shergottite source craters, I need to revisit the issue of how many craters are required.

The number of craters of a particular size striking a particular area over a given period of time can be estimated. I do so by applying the lunar cratering curve to Mars. The lunar flux is a known function and can be adjusted for crater size according to $D^{-2}$, where $D$ is crater diameter (Melosh, 1989). The difference in flux between the moon and Mars is characterized by the factor $R$ (Hartmann, 1999). This flux is multiplied by the surface area of Mars and then by the fraction of Mars of a given age. The total number of impactors is calculated by multiplying this result by the time period of interest, in this case, the CRE age. This is parameterized thusly:

$$N = f_{D=4}f_{D>4}A_{Mars}Rf_{sfc}t_{CRE}$$

(6.1)

where $N$ is the number of impacts, $f_{D=4}$ is the lunar flux for craters 4 km across or larger = $2 \times 10^{-14}$ craters/km²/yr, $f_{D>4}$ adjusts the function for crater size, $A_{Mars}$ is the surface area of Mars = $1.45 \times 10^8$ km², $f_{sfc}$ is the fraction of Mars of a given age, and $t_{CRE}$ is the timespan of interest, in this case determined by the CRE ages of the samples. $R$ is the Hartmann correction factor. $R$ is roughly 1.6, possibly smaller by a factor of 3 or greater by a factor of 2. It is not necessarily a constant with crater size (Hartmann, 1999). Adjusting this function for $D = 3$ km, assuming $N = 4$ impacts (from the CRE data, see Figure 3.2) and setting $t_{CRE} = 4$ Ma, I solve for $f_{sfc}$. The result is 12%. That is, shergottite age material must comprise 12% of the martian surface to be consistent with the geochemical data, the adjusted lunar flux curve, and my own results from modeling impact fragmentation and ejection. The largest uncertainty appears to be in the value for $R$, which would allow values of $f_{sfc}$ for shergottites to range from 6% to 36%. This
result is consistent with the most recent estimates for martian crater ages based on MGS results. As discussed earlier, Tanaka et al. (1992) estimated 2% of Mars as being shergottite age. MGS results will increase that figure by some factor, probably to about 5%. Very young flows are observed on Mars which embay 1 km craters, meaning that the flows are younger than the crater age for the surface in question. Hence the coverage by young flows is underestimated and may be as high as 15% (McEwen, personal communication). Note that this result is based on four source craters for the shergottites, as advocated by Eugster et al. (1997) and Terribilini et al. (1998). In short, my determination of the required crater size for the shergottites is consistent with the constraints of geochemistry and photogeology. Assuming a single impact for the shergottites (Nyquist et al., 1998), the shergottite-aged fraction of Mars would be about 3%, ranging from 1.5% to 9%. These two ranges for the shergottite age fraction of Mars overlap, but only if one is allowed to choose different values of $R$ along with different values of $N$. Besides the difficulties of the single shergottite source crater model discussed previously with regard to compositional (Section 3.1.1) and CRE data (Section 3.2.3) and delivery efficiency (Section 3.5.1), the result appears beyond the admittedly loose bounds of martian photogeology. Hence, I will assume that $N = 4$ impacts for the shergottites.

My best scenario for shergottite launch follows:

Approximately 4 Ma ago, a single event, comparable to or larger in size than a 150 m impactor striking Mars at 10 km/sec, occurred in terrain dominated by 150-200 Ma old lherzolitic material covered by a negligible amount of regolith. On the order of $10^7$ fragments 3-10 cm across were ejected from Mars. Approximately 2% of this material has struck the earth. A comparable amount of material will continue to fall on earth over the next 5 Ma. Three of these fragments fell onto the Antarctic icecap during the last 200 ka. These are known as the lherzolitic shergottites ALH77005, LEW88516, and Y-793605.

About 2.8 Ma ago, a somewhat larger event struck 200-300 Ma old lava flows on or very near the martian surface. Approximately $10^7$ fragments were ejected from
Mars, most of them in the 3-10 cm range, though over $10^6$ were larger than 30 cm. One of these fragments resided on the Antarctic icecap for about 290 ka before discovery. It is known as QUE94201. Two others were observed to fall near the towns of Shergotty, India and Zagami, Nigeria from which they take their names.

Approximately 1.2 Ma ago, another impact into a surface lava flow occurred. The age of this material is uncertain, but may be as large as 800 Ma. The impact was comparable to the basaltic shergottite event. One of the fragments fell over the Dar al Gani section of the Sahara sometime in the last 85 ka. It fragmented in the earth's atmosphere, producing two stones known as DAG 476 and DAG 489.

Finally, approximately 0.8 Ma ago, a similar event struck basaltic material of undetermined age, but probably similar to the other shergottites listed. This stone fell on the Antarctic icecap 170 ka ago and is known as EET79001. Of the shergottite-launching events, it appears this one and the Dar al Gani event are the most likely to yield additional samples.

Debris from all four impacts should still be falling on the earth. The delivery timescale is such that about one-half of earthbound material is delivered in the first 10 Ma years after ejection. Numerical integration shows that 4% of the total number of ejected fragments will hit earth in 10 Ma. Assuming that an impact ejects $\sim 10^8$ fragments, four impact events for the shergottites implies on the order of 10 new martian meteorite falls per year, or approximately one a month. Gladman derived a similar figure based purely on fall statistics and his studies of martian meteorite delivery (Gladman, 1997). A test therefore of my result is to see whether additional samples can be paired with the four impact events as I have defined them. If numerous additional craters are required, then it would indicate that my results are too conservative, or that there is some important process that has been neglected. The rain-of-debris image is supported by considering the relevant timescales. The half-life for earth delivery is 10 Ma, while the wait time for 3 km events on the shergottite-aged part of Mars is roughly 1.5 Ma, with an uncertainty determined by $R$ and $f_{se}$. 
6.2 Impacts into Weathered Material

Older martian meteorites come from more heavily weathered terrain. I reason thus because of the apparent lack of a rock cycle as it occurs on earth. Once formed, rocks are not constantly recycled and reformed as on the Earth. Good evidence for plate tectonics appears to be lacking, despite the most recent MGS results (Malin et al., 1998). If plate tectonics ever occurred on Mars, it was very ancient. Hence, older units are probably more deeply buried by impact-generated regolith.

The growth of regolith has been studied in detail by Shoemaker et al. (1969) with regard to the moon. The depth of regolith is estimated from the size at which the crater population is saturated. The same conceptual approach has been taken with regard to martian regolith by Hartmann (1999). He estimates from the most recent MGS data that Mars is covered globally by 3-5 m of regolith, assuming a 100 Ma age for the youngest martian terrains. Older terrains have deeper regoliths. Hartmann estimates at least of order 10 m for 1 Ga terrain, while the ancient highlands should have a megaregolith a few kilometers deep. Quantifying this runs afoul the same shortcomings as estimating absolute crater ages. The regolith depth is related to the abundance of craters, which has not been tied down to a particular age via radioisotopic studies.

My calculations indicate that 6.7 km diameter craters are probably too small to eject the nakhlites and Chassigny based on their low indicated shock pressures (Table 3.2). I will assume that a 10 km crater (a 650 m diameter projectile at 10 km/sec) is necessary to eject material from ~30-100 m deep regoliths. Using the function above, selecting \( N = 2 \) and \( t_{CI} = 11.5 \) Ma, I estimate that 23% of the martian surface should be nakhlite age. This figure varies from 12% to 69% for different values of \( R \). These figures are roughly twice that for shergottite-aged material. Naturally, the figures are halved if Chassigny and the nakhlites can be paired. Assuming the 23% figure is correct, the wait time for 10 km impacts onto the nakhlite-aged portions of Mars is about 9.4 Ma, comparable to the delivery
halflife. Hence it is unlikely the earth samples material from several impacts into old terrain at once. Rather, the samples are dominated by debris from the most recent event. Note that $f_{sf/e}$ can formally be as large as 69% if $R$ is really the lowest plausible value and Chassigny and the nakhlites cannot be paired. Pairing these stones reduces the upper bound for $f_{sf/e}$ by a factor of two to 35%. However, that value of $R$ would imply a $f_{sf/e}$ for shergottite material of 36%. In other words, one-third of Mars would be shergottite age, one-third would be nakhlite age, and one-third would be older, dating back to heavy bombardment. Such a martian timescale scale appears untenable, providing additional constraints on $R$.

Regarding ALH84001, the required crater may well be larger and the wait time correspondingly longer, given a correlation between regolith depth and terrain age. However, the $f_{sf/e}$ should be larger as well and these two effects could well mitigate each other. Assuming that the required crater is 20 km across (corresponding to a 1600 m diameter impactor at 10 km/sec, the actual number I cannot supply, only to say that it is larger than the nakhlite case) and further assuming that 50% of Mars consists of terrain containing abundant relics from heavy bombardment, applying the lunar flux curve as before results in 1.3 expected candidate impacts for ALH84001.

6.3 Discussion

It appears that the primary factor controlling the efficiency of spall ejection on Mars is the presence and depth of the regolith. I have shown that the presence of a regolith reduces spall velocities, necessitating larger impacts to produce the same result. Moreover, the effect appears to operate over a factor of four in impactor diameter. Extrapolating this effect another factor of two, the implied regolith depth approaches the maximum suspected to exist on Mars. This becomes a key factor in understanding martian meteorite provenance. The principle caveat in this analysis is the small number of known samples.
It has been estimated that approximately 3-5 meters of regolith covers terrain on the order of 100 Ma old. This is why we do not have any $2\pi$ irradiated martian rocks. The regolith covering terrain as old as the youngest martian samples is thick enough to protect the solid rocks from cosmic ray bombardment. Additionally, the solid rocks observed on the surface at each lander site must not get ejected as solid rocks, else we would again have $2\pi$ irradiated samples. It is possible that the coupling between a porous regolith and the solid rocks is poor, and the rocks are not accelerated to escape velocity, or the material in that layer is itself never accelerated to escape velocity. Alternatively, porosity in a thin regolith layer traps so much impact energy as heat that any solid samples melt. This is somewhat speculative until a good model for porosity is developed. Moreover, we are now discussing the fate of individual rock fragments, which is beyond the ability of the SALE2D hydrocode.

Older material probably means deeper regolith. Since the presence of a thick regolith layer can have a profound influence on the minimum size crater to launch martian meteorites, this target dependence is the likely cause of the shergottite age paradox. The younger material is greatly overrepresented in the terrestrial collection because younger material requires smaller, therefore more frequent, impacts to accelerate rocks to Mars' escape velocity. It is possible this target dependence is responsible for the loose correlation between CRE age and crystallization ages of the martian clan meteorites. The rarer the required event, the less likely it occurred recently. That appears to be the case, where the greatest CRE age is for the 4.5 Ga old ALH84001 and all the youngest CRE ages are for the 150-300 Ma shergottites. This argument holds if the crystallization ages for some shergottites is pushed back to 800 Ma, as has been argued for DAG 476. The nakhlites and Chassigny represent intermediate cases. This is to some extent a qualitative argument since the museum efficiency is so very low ($10^{-6}$) and the sample size is rather small (13 and counting). One does not, nor would expect, to see a one-to-one correspondence between CRE and crystallization ages amongst the individual shergottites. However this does explain the broader correlation.
6.4 Conclusion

The martian clan meteorites were launched in 6 or 7 different impact events. The minimum required size of the impact depends on the target composition. Since the target composition can be correlated with terrain age, it follows that it requires a larger impact to liberate material of older vintage. Quantitatively, the relative abundance of shergottites and Chassigny/nakhlites can be explained in this way. Shergottite parent craters need be only 3 km across. Craters 6.7 km across are probably too small to eject solid material if the regolith layer is 30 m or more deep. Assuming a minimum crater of 10 km for Chassigny and the nakhlites one would expect 2 impacts into an area comprising about 23% of Mars. Within the error limits this is not an unreasonable amount. Requiring an even larger crater for ALH84001 (because of a deeper regolith), a single crater would be expected in the last 14 Ma even if 50% of Mars surface dates to heavy bombardment.

I have constructed a scenario for ejecting the martian meteorites that meets all the constraints to the extent that I am aware of and understand them. Clearly, precision is lacking for several key parameters. Therefore, this model is not tightly constrained. However, for the first time, our understanding of impact physics is conjoined to geochemistry, cratering statistics, and celestial mechanics to produce a coherent model of martian meteorite provenance.
CHAPTER 7

Summation

In the preceding six Chapters I have reviewed the major results and many of the nuances regarding the petrology, geochemistry, orbital dynamics of the martian clan meteorites. To this I have added my results from numerical modeling of fragmentation and spallation. Analysis of this large body of work, to which I have contributed but one facet, has resulted in a new model for martian clan meteorite launch and delivery that successfully meets all the observational and theoretical constraints with few exceptions. At this time I will review what has been accomplished as a result of this work and identify areas for further research.

7.1 What has been Resolved

The principal result of this work is a new, reliable insight into the problem of launching meteorites from their parent bodies. The immediate application of that insight is resolving many of the issues that have been argued about the martian clan meteorites for 20 years.

The first paradox was the problem of shergottite launch. Given the original estimate of the minimum required crater (12 km) and the portion of Mars thought to be shergottite age (2%), one would expect 0.04 candidate shergottite impacts. Increasing the portion of Mars thought to be shergottite age to 15% is not helpful—one still expects less than one impact. This paradox is resolved by the discovery that 3 km impacts will suffice. The expected number of impacts is about 4. For the first time, impact physics and photogeology produce numbers that agree to the
same order of magnitude. The precision of these results is still quite poor, owing to the current knowledge about the absolute ages of martian terranes.

The second paradox was the large number of shergottite-producing impacts relative to the number of all other meteorite-launching impacts on Mars. This is resolved by the demonstration that target composition can have a profound influence on shock wave morphology, spall velocity, and energy deposition. If the regolith depth is correlated with terrain age, then the number of expected impacts decreases with terrain age even though young terrains are relatively rare. This explains the observed correlation between martian meteorite CRE age and crystallization age. Older terrains are sampled only by large rare impacts for which the recurrence interval may be comparable to the delivery half-life. Young terrains are sampled much more frequently and Earth sees a steady rain of this material.

7.2 What Remains to be Done

As important and exciting as the above accomplishments are, there remains several areas for further work in the problem of martian meteorite provenance. Aside from the obvious need to populate Mars with teams of geologists armed with rock hammers and mass spectrometers, there is much to be done Earthside.

First of all, because I have used a 2D hydrocode, my computations are necessarily of vertical impacts only. The most probable impact angle is 45 degrees. In oblique impacts the shock wave is no longer axially symmetric with respect to its strength but is stronger downrange and weaker uprange. Examining the importance of this simplification is rather easy given the method used in this study. A 3D Eulerian calculation can be performed, employing Lagrangian tracers as before, but using several sets distributed about the impact site. Each of these sets of tracers would then be used as an input boundary condition for a 2D Lagrangian calculation of fragmentation and spall velocity. This would be efficient in terms of computer time since the 3D calculation would be for a very short duration, and only one would
be required for a suite of second-step 2D calculations. Such a study would better constrain the minimum crater required to eject the martian meteorites. I expect that the results from such a study would adjust the minimum required crater sizes somewhat, but not by more than the uncertainty in applying the lunar flux to Mars. In any case, the adjustment would be consistent for all terrains, so the relative abundance of various source craters would not change.

Increased resolution would make it possible to model better the regolith on young terrains. Presently, the thinnest regolith would be represented by a single layer of cells. It would be most difficult to divine any alteration of the shock wave with this limitation in place. Increased grid resolution would also permit modeling of the larger impacts I think are required for the oldest samples. At 5 m/cell, the 1600 m diameter impactor would have a radius of 160 cells, three times greater than any successful run I completed. Further exploration of the parameter space is always possible.

The importance of layering and regolith properties strongly motivates the development of good constitutive models for this type of material. Ideally we should obtain soil samples that can be experimented upon in the laboratory to determine these models. Closer to home, a breakthrough in the modeling of porosity is required. Such a model would have immediate consequences for this work, and would be applicable to many other problems in planetary science.

At this point, it appears that the most important advances in this line of research will come from additional sample acquisition and establishing an absolute dating scheme for the martian surface.
REFERENCES


The Bhaktivedanta Book Trust.


