Angelo Pio Rossi
Stephan van Gasselt
Editors

Planetary Geology

Springer

PRAXIS

Planetary Geology

Springer Praxis Books

Astronomy and Planetary Sciences

Angelo Pio Rossi • Stephan van Gasselt
Editors

Planetary Geology

Springer

PRAXIS

Editors
Angelo Pio Rossi
Jacobs University Bremen
Bremen, Germany

Stephan van Gasselt
National Chengchi University
Taipei, Taiwan

Springer Praxis Books
ISSN 2366-0082 ISSN 2366-0090 (electronic)
Astronomy and Planetary Sciences
ISBN 978-3-319-65177-4 ISBN 978-3-319-65179-8 (eBook)
DOI 10.1007/978-3-319-65179-8

Library of Congress Control Number: 2017957712

Cover illustration: Mosaic of 20 images (AS17-140-21488 to 21507) Apollo 17, Station 6. Credit: All Imagery NASA / Public Domain
Cover design: Jim Wilkie

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To our family and friends. For our students.

Foreword

Planetary geology is rapidly becoming one of the most exciting fields of science as our species increases our capability of studying solar system objects. More missions, carried out by a larger number of nations, have provided vast amounts of data, most of which is publically available. It is not surprising that this treasure trove has attracted interest from geologists like me, who were trained and worked on terrestrial geology. While making such a transition can be rewarding, it also comes at a cost. Despite having taught portions of an introductory astronomy course for years, I found that the switch from Terrestrial Proterozoic to Martian Hesperian problems required a steep learning curve. While the geological problems were similar, much of the background, the data and the methods used were very different than those that I had been used to. I made the transition to studying Martian geology by reading background material and with the help of patient explanations from kind collaborators, several of whom have contributed to this book. I believe this book would have been very useful to me during this transition and I think it will be a useful resource in the upcoming years.

The obvious strengths of this book are that it covers a broad range of topics in planetary sciences and that it does not require a background in a specific scientific discipline. This is particularly important because the appeal of planetary geology attracts science students from diverse backgrounds. Students with Biology or Physics backgrounds will benefit from the explanation of basic geological concepts such as the principles of stratigraphy. Most science students know very little about image processing or map projections. I believe that this book is ideally suited for a Planetary Geology course at a senior undergraduate level that is open to students who are not traditional geology majors and I plan to adopt it as a text for just such a course. The topics covered in this book will allow me to cover common ground whilst enabling students to tailor aspects of the course to their particular interests. I expect that student projects based on the frontiers section of the book will be popular.

Because of the constant stream of data and imagery provided by numerous missions, it would be nearly impossible for any book to represent the latest understanding of any given topic. The strength of this book is its focus on basic

principles and the role of geological processes in shaping planetary surfaces. Geological principles are nearly timeless in that they do not change often; the fundamental shift that resulted from understanding plate tectonics took place more than 50 years ago. This is, after all, why I am still able to apply concepts learned during the later parts of a previous millennium to the study of another planet today. That is why I believe this book is an excellent introduction and overview of planetary science and it will continue to be useful for the foreseeable future.

St. Catherines, Canada Frank Fueten
May 2017

Twenty years ago Mars was a different planet. During that time, the Magellan mission to Venus provided first insights into incredibly exotic environments. The icy bodies of the Outer Solar System were still largely unknown—apart from the magnificent data return of the Voyager missions. Mars represents a compelling example of the recent evolution of a geological perspective. According to the common view of the (pre)-Viking era Mars was a planet dominated by volcanic processes, and sedimentary activity was limited to *recent* aeolian processes. A few unsung authors published papers or maps suggesting the past existence of lakes, rivers, fan deltas, alluvial fans, but it took a considerable time to widely realise that water has been one of the main agents (if not the major agent) in shaping the surfaces of both terrestrial and icy bodies with an atmosphere. Nowadays Mars is seen as a planet with a complex history. Its surface has been carved by rivers, lakes, glaciers, deltas: in a few words by the entire sequel of continental and shallow water processes. Stephen Jay Gould, the great and popular Agassiz Professor of Zoology at Harvard—who actually was a palaeontologist—used to say that one is able to find what one already knows.

Titan shares some similarities. The pre-Cassini view of the mysterious moon of Saturn was that of something as remote as possible from Earth: just a single ocean or a little continent. It was seemingly hard to depict a planet in Terrestrial terms as a planet was compulsorily something different from Earth. Instead it would have been rather possible to envisage a planet with flowing liquids and the formation of lakes or seas with extensive and flat shorelines, dunes and ergs. Sedimentary geology was not that popular in these days.

This book collects the fresh and frank views of a group of mostly young (at least, young for someone of my age) and energetic scientists. They deliver the right sequence of subjects to the reader starting from the conceptual and practical tools for the geological analysis of planets, going through processes and closing with present frontiers. The reader is accompanied, taken by the hand, throughout the book and its topics. Students will find the right balance of topics developed as a (field) trip starting from the fundamentals of geology and ending with geologists on Mars. In addition they will be able to continuously relay with their Terrestrial background, comparing different settings, features and processes among the planetary bodies.

Sedimentology, sedimentary geology and the sedimentary record are pervasive in this book. Of course, they apply mostly to planets with an atmosphere, although

some sedimentary processes may occur on the Moon, asteroids, icy satellites and comets. Like on Earth we can reconstruct the paleoclimate evolution by looking at the type of environments and sediments that have been present at the surface. Outcrops can now be imaged in 3D and stratal patterns and details of stratifications can be studied. The sedimentary records are remarkable archives of their environmental and climate change.

The search for life is also concentrated on sediments that are the best geological objects to harbour evidences of its presence and planetary sedimentary records are good targets for investigation. Among many other aspects, this book provides to both students of geological sciences and researcher of planetary sciences a sober account of this new view of planetary geology.

Marrakech, Morocco Gian Gabriele Ori
June 2017

Preface

Many excellent books on planetary geological topics exist, especially at graduate levels, and they are complemented by a number of carefully written introductory textbooks on planetary science.

The present textbook has been designed to cover the gap in between, presenting an up-to-date view written by authors covering their own fields of research. The book has been edited to provide an integrated overview of the diverse components of current Planetary Geology.

Planetary Geology as a professional discipline is practiced by Earth scientists and enriched by researchers and teachers with backgrounds in topics of physics, chemistry, astrophysics or biology. A geological background is not necessarily present in these study areas, nor do traditional geoscience programmes usually include in-depth specific topics.

This book can be in principle used by any of the categories above, provided with some basics of geoscience, both for general education and information and for easing the perspective of undergraduate to early graduate research efforts.

The primary audience consists of advanced undergraduate geoscience students, as well as early graduate non-geoscience students in space-related topics.

The book would serve as a manual to approach Planetary Geology from its basic tools and methods, going through the main geological processes acting on Solar System bodies. Each chapter points towards deeper or more specific sources of information. Up-to-date practical aspects are introduced in the appendix.

The book is divided into 4 parts: Methods and tools (Chaps. 1–5), Processes and sources (Chaps. 6–10), Integration and geological syntheses (Chaps. 11–13), Frontiers (Chaps. 14–15).

Chapter 1 (A.P. Rossi and S. van Gasselt) provides a brief overview of general concepts, including some historical aspects. Chapter 2 (M. Pondrelli et al.) summarises the intellectual tools used in geology and applied to planetary science. Chapter 3 (S. van Gasselt et al.) introduces the various techniques and approaches used in planetary geological exploration. Chapter 4 (T. Hare et al.) covers the cartographic and geological mapping background of planetary geology, while Chap. 5 (A.P. Rossi and S. van Gasselt) covers its ground truth via either human

or robotic explorers. Chapter 6 (A. Greshake and J. Fritz) introduces meteorites. Chapter 7 (T. Kenkmann and G. Wulf) describes the basics of impact cratering and its role in the various Solar System bodies. Chapters 8 (E. Hauber et al.) and 9 (N. Mangold) provide a view respectively on interior and surface processes active on celestial body. Chapters 11 (A.P. Rossi et al.), 12 (R. Wagner et al.) and 13 (N. Schmedemann et al.) integrate concepts and methods introduced in previous chapters for respectively terrestrial planets, icy satellites and small bodies, including dwarf planets. Finally, two topics of importance for future exploration are introduced: Chap. 14 (B. Cavalazzi et al.) describes astrobiological topics related to planetary exploration, while Chap. 15 (A. Abbud-Madrid) introduces the emerging field of planetary resource geology.

The scope and extent of the treatment of each topic is necessarily limited and aims at a general understanding, particularly from the methodological point of view. We hope our readers will develop a feeling for planetary geoscience, allowing them to eventually dig further.

For detailed information on each specific Solar System target, suggested further readings are indicated for each chapter. When possible, those readings are books or review papers. For very recent developments, suitable research papers are suggested.

The nomenclature used in this book is respecting current use in the Planetary Science community and naming conventions of Solar System bodies and their toponyms respect the IAU rules. As an example, the current mention of Pluto is as *Dwarf Planet* rather than planet, a process which occurred in the last decade.

In the case of landers and rovers, and to some extent for recent orbital or fly-by missions, the use of informal names even in the scientific literature is getting more common. Reference to informal names is avoided, although in some case (Chap. 5) mentions of this sort are present and highlighted as informal, due to the relatively high-pace nature of rover exploration, the small size of features and the need for naming them effectively by experiment teams.

At the time we started this editorial project, we had some plans, expectations and hopes on the outcome. There have been obviously difficulties in putting together the product we designed and desired, largely related to tasks and professional duties of us all. *It always takes longer*: that also applied to the present book.

We had the honour and luck to have a team of motivated authors that made the exercise if not easier, at least interesting and pleasant, with all the natural delays and small change of plans embedded in such type of effort, particularly when trying to combine specific topics and deliver to a broad, not specialised audience.

Probably much is going to change in just few years time in our understanding of the overall geological evolution in the Solar System. The material presented throughout the book attempts to capture the state of the art after about five decades of modern Planetary Geology, and some five centuries or so of Planetary *geological* astronomy.

We hope to have achieved at least part of what we had aimed for and that the core target of such book is served by our collective work: Students already engaged in geologic training but not much exposed to the exotic yet familiar extraterrestrial geology.

We also hope that, beyond academic use, some of the materials in this book might be useful to point to the vast amount of freely accessible resources describing or depicting the geology of our Solar System. There is a growing amount of openly accessible data and information, as pointed out in the appendix and links therein. Planetary geology *virtual hikes* are not for professional scientists or topical experts only, they are there for all curious individuals. Latest technological development in data handling and visualisation do and—even more—will allow citizen planetary science (self) education. In our hope, some of the potential readers of this book might be helped in such endeavour.

Bremen, Germany — Angelo Pio Rossi
Berlin, Germany — Stephan van Gasselt
June 2017

Acknowledgements

The editors are indebted to Johannes Geiss, Roger-Maurice Bonnet and Martin Huber for their encouragement in pursuing our idea of an accessible textbook on Planetary Geology.

Johannes Geiss in particular deserves special thanks for having shared with unique humour incredible first-hand experiences of the birth of modern Planetary Science, from Apollo times onwards.

Our editorial gratitude goes also to all those colleagues who kept us motivated throughout the years, few of which embarked in the endeavour resulting in the present book. Their kind presence, ranging from scientific cooperation, kind encouraging words or just silent friendly support, is most appreciated: Even if not specifically listed, they know how much their support counts for us.

We thank also those who exposed (or where exposed together with us) us to Planetary Geology for the first time, during late undergraduate, or early graduate times. Working a running planetary mission, ESA Mars Express, from the mid 2000s has been very valuable for both editors. We wish to thank also all the people we interacted during those times, across Europe and beyond, as well colleagues with whom we shared scientific, educational or community support duties in all these years. An inspiring and motivating experience was the e2e iSAG from the early 2010s: a figure out of the report made it here, so the memory of insightful conversations with all of its members.

The International Space Science Institute (ISSI) in Bern (Switzerland) as a whole has been helpful during the early phases of the book conception, as well as ISSI Teams, Workshops and Working Groups throughout these years provided additional motivation and inspiration.

The students we met through these years provided most motivation to the editors. We hope they could find something useful out of this book, regardless of their career path.

We are extremely grateful to Frank Fueten for his generous and patient review of the geological content of our book.

Our thanks also go to Laetitia le Deit, Jessica Flahaut, Mikhail Minin and Ramiro Marco Figuera.

We are grateful to Ramon Khanna for considering and supporting our contribution to the Springer series and to Alessia Valdarno for her patience and support.

APR wishes to offer belated thanks to those inspired him, warning about the risks of academy yet encouraging towards the study and beauty of Geology. His path brought him to work on topics very far from those taught a couple of decades ago by Giovanni *Jack* Pallini.

Last but not least, APR warmly thanks his family and its patience in accommodating editing and writing sessions off-working hours, as well as his kids for painting and adding stickers mostly on the backside of his laptop screen.

We tried to limit mistakes and detect them as they occurred. Some might have remained and some accidentally added during the editorial work. Such errors should be considered solely the responsibility of the two editors.

Chapter-specific acknowledgements are included separately.

Contents

Contributors

Angel Abbud-Madrid Center for Space Resources, Colorado School of Mines, Golden, CO, USA

Victor R. Baker Department of Hydrology and Water Resources, School of Earth and Environmental Sciences, The University of Arizona, Tucson, AZ, USA

André Brack Centre de Biophysique Moléeculaire, Orléans Cedex 2, France

Doris Breuer Institute of Planetary Research, Experimental Planetary Physics, German Aerospace Center, Berlin, Germany

Petr Brož Institute of Geophysics, Czech Academy of Science (ASCR), Prague, Czech Republic

Sherry Cady Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA

Barbara Cavalazzi Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, Bologna, Italy

Mario d'Amore Institute of Planetary Research, Experimental Planetary Physics, German Aerospace Center, Berlin, Germany

Joerg Fritz Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Museum für Naturkunde, Berlin, Germany

Mihaela Glamoclija Department of Earth and Environmental Sciences, Rutgers University - Newark, Newark, NJ, USA

Ansgar Greshake Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Museum für Naturkunde, Berlin, Germany

Trent M. Hare Astrogeology Science Center, United States Geological Survey, Flagstaff, AZ, USA

Ernst Hauber Institut für Planetenforschung/Institute of Planetary Research, Planetary Geology, German Aerospace Center (DLR), Berlin, Germany

Harald Hiesinger Westfälische Wilhelms-Universität Münster, Münster, Germany

Thomas Kenkmann Institute of Earth and Environmental Sciences-Geology, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany

Randolph L. Kirk Astrogeology Science Center, United States Geological Survey, Flagstaff, AZ, USA

Damien Loizeau Université Lyon 1, Villeurbanne, France

Nicolas Mangold Lab. de Planétologie et Géodynamique de Nantes, CNRS et Université de Nantes, Nantes, France

Matteo Massironi Dipartimento di Geoscienze, Università di Padova, Padova, Italy

Daniel Mège Space Research Centre, Polish Academy of Sciences, Warsaw, Poland

Roberto Orosei Osservatorio di Radioastronomia, Istituto Nazionale di Astrofisica, Bologna, Italy

Thomas Platz Max Planck Institute for Solar System Research, Göttingen, Germany

Monica Pondrelli International Research School of Planetary Sciences, Università d'Annunzio, Pescara, Italy

Angelo Pio Rossi Department of Physics and Earth Sciences, Jacobs University Bremen, Bremen, Germany

Nico Schmedemann Department of Earth Sciences, Institute of Geological Sciences, Freie Universität Berlin, Berlin, Germany

James A. Skinner, Jr. Astrogeology Science Center, United States Geological Survey, Flagstaff, AZ, USA

Katrin Stephan Institut für Planetenforschung/Institute of Planetary Research, German Aerospace Center (DLR), Berlin, Germany

Nicola Tosi Institute of Planetary Research, Experimental Planetary Physics, German Aerospace Center, Berlin, Germany

Stephan van Gasselt National Chengchi University, Taipei, Taiwan

Roland Wagner Institute of Planetary Research, Experimental Planetary Physics, German Aerospace Center, Berlin, Germany

Frances Westall Centre de Biophysique Moléeculaire, Orléans Cedex 2, France

Gerwin Wulf Institute of Earth and Environmental Sciences-Geology, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany

Acronyms

ALSE	Apollo Lunar Sounding Experiment
AOA	Amoeboid olivine aggregates
APT	Adenosine triphosphate, coenzyme transporting chemical energy in organisms' cells
APXS	Alpha Particle X-ray Spectrometer, experiment, e.g. on MER, MSL
ASP	AMES Stereo Pipeline, digital stereogrammetry software package
ASTM	American Society for Testing and Materials
AU	Astronomical Unit, mean Sun–Earth distance, equivalent to about 150 million *km*
BIF	Banded Iron Formation
CAI	Calcium–Aluminum-rich inclusions
CCAM	Carbonaceous Chondrite Anhydrous Mineral
CCD	Charge-Coupled Devices
CDR	Calibrated Data Record
CHIMRA	Collection and Handling for Interior Martian Rock Analysis, on board MSL
CIVA	Comet Infrared and Visible Analyser, experiment on Philae, ESA Rosetta lander
CMB	Core–Mantle Boundary
CMOS	Complementary Metal Oxide Semiconductor
CNSA	China National Space Administration
CODMAC	Committee on Data Management and Computation
CONSERT	COmet Nucleus Sounding Experiment by Radiowave Transmission
COSPAR	COmmittee on SPAce Research
CRISM	Compact Reconnaissance Imaging Spectrometer for Mars, MRO experiment
CRM	Chemical remanent magnetisation
CSDGM	Content Standard for Digital Geospatial Metadata
CSFD	Crater Size Frequency Distribution

CTX	Context Camera, NASA MRO experiment
DDR	Derived Data Records
DEM	Digital Elevation Model
DiCE	Dissected Crater Ejecta
DISR	Descent Imager/Spectral Radiometer, experiment on board Huygens probe, on Cassini–Huygens
DMA	(United States') Defense Mapping Agency
DN	Digital Number
DRT	Dust Removal Tool, on board MSL
DTM	Digital Terrain Model
EDR	Experiment Data Record
ESA	European Space Agency
ESAC	European Space Astronomy Centre, ESA establishment
ETM	Enhanced Thematic Mapper, Landsat 7–8
EVA	Extravehicular activity
FGDC	Federal Geographic Data Committee
GEO	Geosynchronous equatorial orbit
GIS	Geographic Information System
GOE	Great Oxidation Event, around 2.5 Gyr ago on Earth
GPR	Ground Penetrating Radar
GPS	Global Positioning System
GRS	Gamma Ray Spectrometer, e.g. a NASA ODY experiment
HED	Howardites, eucrites and diogenites meteorite group
HiRISE	High Resolution Imagine Science Experiment, NASA MRO experiment
HRSC	High Resolution Stereo Camera, ESA MEX experiment
IAU	International Astronomical Union
IDP	Interplanetary dust particle
InSight	Interior Exploration using Seismic Investigations, Geodesy and Heat Transport, NASA mission
IR	Infrared, portion of the electromagnetic spectrum
ISIS	Integrated Software for Imagers and Spectrometers, developed by USGS
ISO	International Organization for Standardization
ISRO	Indian Space Research Organisation
ISRU	In-Situ Resource Utilisation
ISSI	International Space Science Institute, research institute in Bern, Switzerland
IUGS	International Union of Geological Sciences
JAXA	Japan Aerospace Exploration Agency
JPL	Jet Propulsion Laboratory
KBO	Kuiper Belt Object
KREEP	Potassium (K), Rare Earth Elements (REE), Phosphorous (P), Moon terrane characterised by incompatible elements
LASER	Light Amplification by Stimulated Emission of Radiation

LCROSS	Lunar Crater Observation and Sensing Satellite, NASA Mission
LEND	Lunar Exploration Neutron Detector, LRO experiment
LEO	Low Earth orbit
LHB	Late Heavy Bombardment
LIBS	Laser Induced Breakdown Spectrometry
LIDAR	Light Detection and Ranging
LIP	Large Igneous Province
LLD	Light-toned Layered Deposits
LM	Landing Module
LOLA	Lunar Orbiter Laser Altimeter, NASA LRO experiment
LRCROSS	Lunar CRater Observation and Sensing Satellite
LRO	Lunar Reconnaissance Orbiter, NASA mission
LRO	LRO Camera, experiment on board NASA LRO
LRR	Laser Ranging Retroreflector, Apollo experiment
M3	Moon Mineralogy Mapper, on board ISRO Chandrayaan-1
MAHLI	MArs Hand Lens Imager, NASA MSL experiment
MARA	MAscot RAdiometer, on board DLR MASCOT
MARSIS	Mars Advanced Radar for Subsurface and Ionosphere Sounding, ESA MEX experiment
MASCAM	MAScot CAMera, on board DLR MASCOT
MASCOT	Mobile Asteroid Surface Scout, DLR lander on board JAXA Hayabusa-2
MASMAG	MAScot MAGnetometer, on board DLR MASCOT
MDIS	Mercury Dual Imaging System, NASA Messenger experiment
MER	Mars Exploration Rovers, NASA mission with 2 rovers
MERTIS	MErcury Radiometer and Thermal Infrared Spectrometer
MESSENGER	(MErcury Surface, Space ENvironment, GEochemistry, and Ranging, NASA Mission
MEX	Mars Express, ESA mission
MGS	Mars Global Surveyor, NASA mission
Mini-TES	Miniature Thermal Emission Spectrometer, NASA MER experiment
MIR	Middle-wave infrared
ML	Mobile lid
MLA	Mercury Laser Altimeter, NASA Messenger experiment
MMR	Mean Motion Resonances
MOLA	Mars Orbiter Laser Altimeter, NASA MGS experiment
MOMA	Mars Organics Molecule Analyser, ESA ExoMars experiment
MRO	Mars Reconnaissance Orbiter, NASA mission
MSL	Mars Science Laboratory, also known as Curiosity, NASA rover
MSR	Mars Sample Return
MSS	Multispectral Scanner, Landsat 1–3
NAIF	Navigation and Ancillary Information Facility
NASA	National Aeronautics and Space Administration
NEA	Near-Earth Asteroid

NEO	Near-Earth Object
NRM	Natural Remanent Magnetisation
NSSCDA	NASA Space Science Coordinated Data Archive
ODY	Mars Odyssey, NASA mission
OSIRIS-REx	Origins, Spectral Interpretation, Resource Identification, Security, Regolith Explorer, NASA mission to an asteroid, including sample return
PDS	Planetary Data System, planetary data archiving standard, repositories, e.g. PDS3, PDS4 versions
PICS	Planetary Image Cartography System
PILOT	Precursor ISRU Lunar Oxygen Testbed
PKT	Procellarum KREEP Terrane
PSA	Planetary Science Archive, ESA planetary data archive
PSEP	Passive Seismic Experiment Package, Apollo experiment
RAC	Robotic Arm Camera, Phoenix experiment
RADAR	Radio Detection and Ranging
RAT	Rock Abrasion Tool, MER tool
RDR	Reduced Data Record
RNA	Ribonucleic acid, molecule involved in gene codification and decodification
ROSINA	Rosetta Orbiter Spectrometer for Ion and Neutral Analysis
ROV	Remotely Operated (underwater) Vehicle
RPM	Resource Prospector Mission, NASA Lunar mission
RS	Remote Sensing
RSL	Recurrent Slope Linae
RTG	Radioactive Thermal Generator, type of electric power generator for deep space probes and landers/rovers
RWGS	Revers Water Gas Shift, Oxygen production technology
SAR	Synthetic Aperture Radar
SD2	Sample Drilling and Distribution, on board ESA Rosetta
SELENE	Selenological and Engineering Explorer, Kaguya Japanese Moon mission
SESAME	Surface Electric Sounding and Acoustic Monitoring Experiment, on board Philae, ESA Rosetta lander
SFD	Size-frequency distribution, of asteroids
SHARAD	Shallow Subsurface Radar, NASA MRO experiment
SIR	Spaceborne Imaging Radar, different experiments on board several Space Shuttle missions
SL	Stagnant lid
SMART-1	Small Missions for Advanced Research in Technology-1, ESA Moon technology demonstrator mission
SNC	Shergottites Nakhlites Chassignites, Mars meteorites
SOE	Solid Oxide Electrolysis
SPICE	Spacecraft, Planet, Instrument, C-matrix (pointing), and Events
SRM	shock remanent magnetisation

SRTM	Shuttle Radar Topography Mission
SWC	Solar Wind Collector, Apollo experiment
SWIR	Short-wave infrared
TES	Thermal Emission Spectrometer, on board NASA MGS
TGO	Trace Gas Orbiter, ESA ExoMars mission
TIR	Thermal infrared
TM	Thematic Mapper, Landsat 4–5
TRM	Thermoremanent magnetisation
UAV	Unmanned Aerial Vehicle
USGS	United States Geological Survey
UV	Ultraviolet
VESPA	Virtual European Solar and Planetary Access
VEX	Venus Express, ESA mission
VICAR	Video Image Communication and Retrieval
VIRTIS	Visual IR Thermal Imaging Spectrometer, Rosetta, VEX experiments
VIS	Visible
VNIR	Visible and Near Infrared
VO	Virtual Observatory
YORP	Yarkovsky–O'Keefe–Radzievskii–Paddack effect

List of Figures

List of Tables

Part I
Methods and Tools

Chapter 1
Introduction

Angelo Pio Rossi and Stephan van Gasselt

1.1 Planetary Geology as a Discipline

Planetary Geology has evolved to a broad and interdisciplinary research field which is based on terrestrial geologic approaches and which has borrowed expertise, intellectual and analytical tools, as well as technologies from astronomy and astrophysics (Fig. 1.1). Its foundation and principles are still geology although it has to heavily rely on remote sensing rather than fieldwork for obvious reasons. It is very much observational and in many respect not as much experimental as historical geology might be. But it does embed very experimental sub-disciplines and aspects when it comes to experimental petrology, impact cratering, and many other areas. A large fraction of experiments in impact cratering and planetary paloclimate modeling are numerical in nature.

In these decades, the fields of planetary science and geology in particular have captured the attention of both academic, education communities and the general public at large notwithstanding also considerable struggle to substantialise broadly in all modern education systems, mainly caused by fear of contact. The triumph of planetary geology, however, is most likely linked to spectacular images of alien landscapes, to the mystery around far, unaccessible worlds and the possibility that some of them harbour—or might have harboured—life. It is certainly also linked to the overall fascination of the exploration of space and the Solar System, shared also by so many remote and unaccessible and alien-feeling locations on Earth, such as

A.P. Rossi (✉)
Jacobs University Bremen, Campus Ring 1, 29795 Bremen, Germany
e-mail: an.rossi@jacobs-university.de

S. van Gasselt
National Chengchi University, No. 64, Sec 2, ZhiNan Rd., Wenshan District, Taipei 11605, Taiwan
e-mail: svg@nccu.edu.tw

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_1

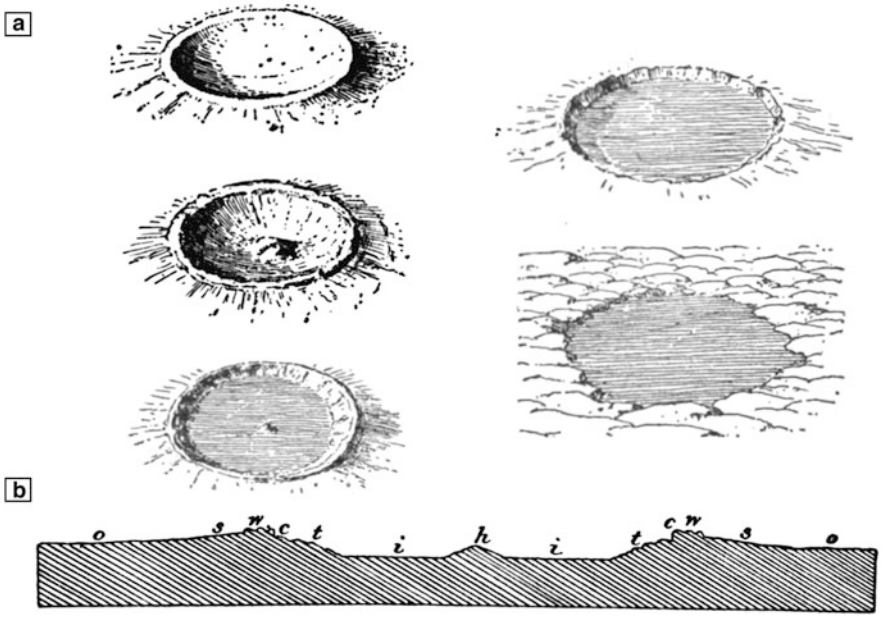

Fig. 1.1 Excerpt from G.K. Gilbert's drawings of craters on the Moon: (**a**) Some key characteristics of simple and complex impact craters (see Chap. 7) are visible from the late nineteenth century drawings. (**b**) Topographical cross-section across a complex crater, highlighting its central peak. Source: Gilbert, 1893

the depth of the oceans or the most inaccessible parts of deserts or mountain chains. The propinquity between extraterrestrial and submarine remoteness is partially also methodological in nature, e.g., handling of marine bathymetry data is not that much different from handling extraterrestrial topography data (e.g., Chaps. 3 and 5).

Beyond that, significant technological advances occurring roughly every 10 years not only kept the fascination alive but also provided the basis for significant leaps in the geological analytical repertoire and scientific understanding. Such milestones comprise first-light observations of planets in the 1960s (Fig. 1.2), detailed orbiter imagery in the 1970s, lander photography in the 1980, high-resolution imaging and detailed topography in the late 1990s, high-resolution spectroscopy and subsurface investigations in the 2000s and the advent of in-situ laboratories in the 2010s.

In matter of just a few decades, point-like astronomical objects were exposed as completely new geological worlds, where our Earth-bound knowledge of processes has sometimes hard times in figuring out their functioning, either for the remote distance in space and time that separates us from them, or because of the deep physical differences that makes their geology so exotic. The recent exploration of the Pluto system (Chap. 13), like those of Jupiter and Saturn (Chap. 12), revealed a mixture of very familiar and extremely surprising features, just when lithospheres are substituted by cryospheres.

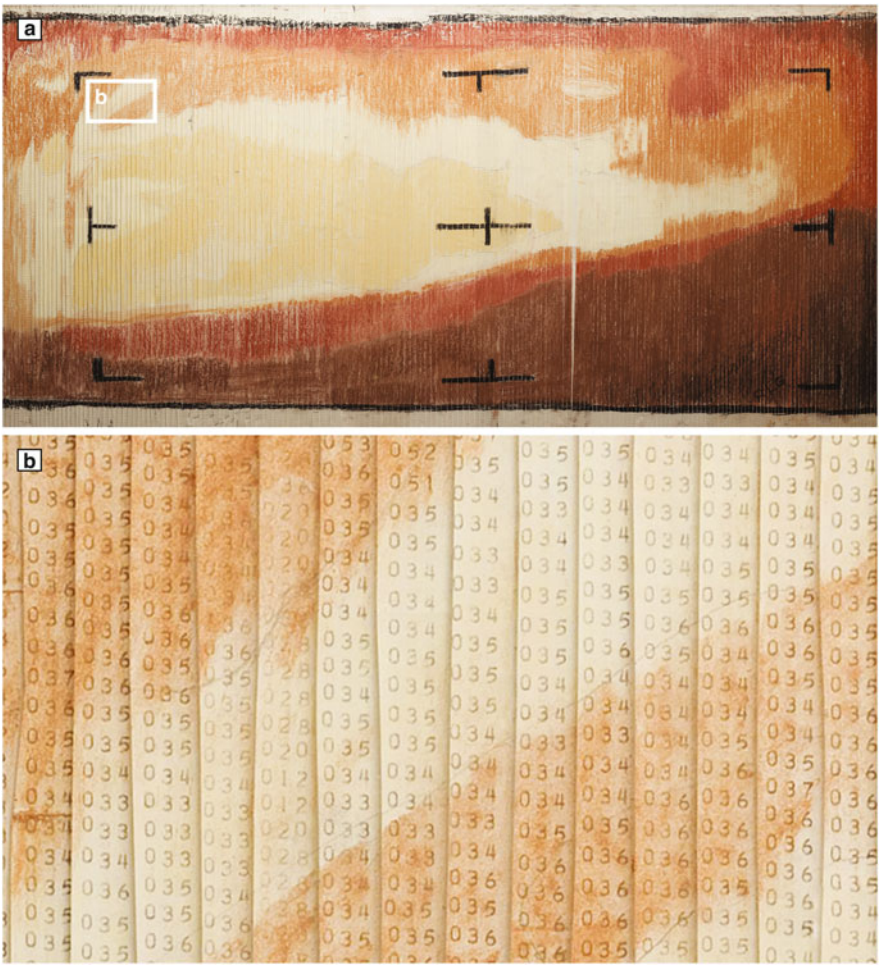

Fig. 1.2 First digital image collected by a spacecraft on Mars, by Mariner 4 in 1965. (**a**) Reproduction of the original one at JPL, hand-drawn by engineers based on received data, the outline of sub-figure (**b**) is indicated in *white*; (**b**) Detail of the hand-drawn digital number classification. The colorisation is only based on DN thresholds. Source: NASA/JPL

The birth of Planetary Geology as a discipline can be traced back either to late planetary astronomical observations or to the early work on impact cratering and related disputes on the nature of craters on the Moon in the last few centuries. It was obviously difficult to establish a direct link to actual geological processes and early observations of very distant planetary bodies, such as Jupiter's Galilean moons, when they were first detected in the early seventeenth century. Consequently, they were only known as astronomical objects.

Planetary Geology is very broad and several disciplines concur to define it, ranging from geology to geochemistry, biology and remote sensing, in addition to

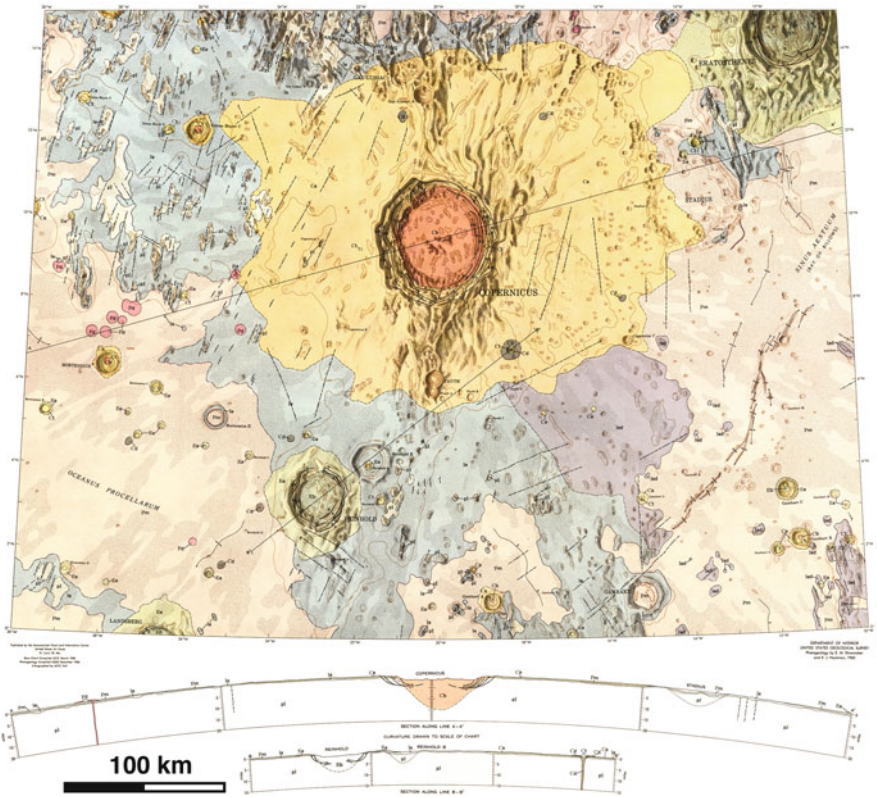

Fig. 1.3 Prototypal geological map of Copernicus Crater by E. Shoemaker in 1960. Although not the first moon map (a global physiographic one was published in 1960 by Hackman and Mason), it is the first geological one to serve as a base for following systematic mapping (Chap. 4). Source: USGS/LPI, P. Spudis

a wide variety of sample science, especially on in-situ and sample return missions (Chap. 3). Neighboring disciplines, which can also be considered part of planetary geology include, among others, cosmochemistry, petrology, (exo)biology, solid Earth geophysics.

It is not by chance that the Moon has been the first planetary geological target (Fig. 1.3), being the closest extraterrestrial object to Earth, as well as the key planetary body that gave birth to Planetary Geology as a modern discipline. Actual geological observations, rather than astronomical ones, are for example those performed by Galileo in Italy of the near side of the Moon (Fig. 1.4). The discipline corresponding to planetary remote sensing would have then been defined as *selenography*.

For this volume, we tried to include broadly accepted views on the different facets of the book, withing and through its chapters. This does not mean all scientists agree on everything that is written here. Interpretations might change based on newer, better data. Themselves, data can be better calibrated or errors detected, that

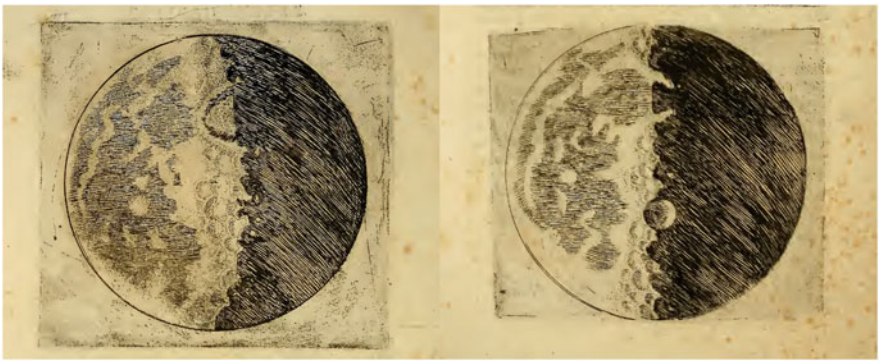

Fig. 1.4 Galileo's drawings of the Moon from the *Sidereus Nuncius*. Major physiographic distinctions are well recognised, including maria as well as large impact basins. Source: reproduced from Sidereus Nuncius, Galilei, 1610

might re-visit earlier interpretations and trigger new ideas. The reader should be aware, though, that very recent results, first-hand exploited by a limited number of scientists, might be more suitable to changes or revision (sometimes, retractions) than others. Please refer to both the reading lists provided at the end of each chapter, and to critical monitoring of the scientific literature.

1.2 The Playground for Planetary Geology

Geology is solid-surface research and this reduces the number of objects to be studied in our Solar System slightly, although not drastically. As of today, the Solar System consists of eight planets, five dwarf planets, 176 moons (or natural planetary satellites) and—as of mid 2016—659,212 known asteroids and 3296 cometary objects with numbers going up at a rate of ten new asteroid detections per day. Except for the giant gas planets Jupiter and Saturn, and the sub-giants Uranus and Neptune, all objects can be studied geologically through remote sensing or in-situ operations including rock sampling (see Chap. 3). Thus far, mankind put landers on the surface of only two other planets beyond Earth (Venus and Mars), two moons (the Earth's moon and Saturn's moon Titan), one asteroid (25143 Itokawa)[1] and two comets (67P/Churyumov-Gerasimenko and 9P/Tempel 1). Classical geologic field studies have only been conducted on the Moon on which two out of three astronauts of Apollo 11–12, 14–17 had the privilege to set their feet.

With resolution 5A of the *26th International Astronomical Union* in Prague in 2006 we have obtained a number of definitions which allow us to distinguish

[1] In September 2016 OSIRIS-REx was sent on its way to reach carbonaceous asteroid 101955 Bennu in 2019.

between different objects in the Solar System, mainly planets and their satellites and the range of small objects that have been found since the last decades with the technological advances of observations programs. The nomenclature on which was agreed by the IAU during that meeting therefore reflect the current understanding that mankind has.

Resolution 5A verbatim states:

1. A *planet* is a celestial body that (a) is in orbit around the Sun, (b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape, and (c) has cleared the neighbourhood around its orbit.
2. A *dwarf planet* is a celestial body that
   a. is in orbit around the Sun,
   b. has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape.
   c. has not cleared the neighbourhood around its orbit, and
   d. is not a satellite.
3. All other objects, except satellites, orbiting the Sun shall be referred to collectively as *Small Solar-System Bodies*.

And with Resolution 6A Pluto became a *dwarf planet* and has been recognized as the prototype of a new category of *trans-Neptunian objects*. That, however was not the first time, the terminology of the Solar System was altered and it probably won't be the last. When Ceres—now a dwarf planet, but an asteroid before 2006—was discovered in 1899, it was considered as the lost tenth planet.

The perception of the layout of the Solar System has faced many changes and they document the development of religious belief and scientific understanding mankind has gone through in history.

Until the beginning of the seventeenth century—during the time of the European Renaissance, the Solar System's layout has been considered to be geocentric for almost 2000 years, although early Greek astronomers brought up an early concept of a heliocentric layout about 300 BC. Among the many advocates of a geocentric layout, famous names of astronomers and philosophers such as ARISTOTLE (384–322 BC) or CLAUDIUS PTOLEMY (100–170 AD) can be read. The latter gave name to the Ptolemy system, a very special geocentric system in which planets and the Sun move along epicycles around the Earth in order to describe some of the observations which could not be explained by a simple movement along circles.

When NICOLAUS COPERNICUS (1473–1543) published his work on heliocentrism shortly before his death, the stage was set for the *Scientific Revolution* and despite attempts to eliminate his theories, scientific observations during the upcoming centuries provided further observational evidence in favour of a heliocentric system. Although TYCHE BRAHE (1546–1601) tried to combine aspects of the geocentric Ptolomean system and the heliocentric Copernican system and published his work based on detailed observations (later known as Tychonic system), it was his student JOHANNES KEPLER (1571–1630) who finally succeeded to describe the

motion of planets within a heliocentric system. This success, however, was only possible with the help of Brahe's observations. Kepler's first two laws of planetary motion were published in *Astronomia Nova* and the third law in *Harmonices Mundi* about 10 years later. This work heavily relied on observations by telescopes—a tool that was invented and further refined in 1610 by GALILEO GALILEI (1564–1642) and by observations of the large satellites of Jupiter—later named the *Galilean satellites*.

It was years later that ISAAC NEWTON (1642–1726) provided the actual mathematical foundations of celestial mechanics with which Kepler's laws could be described and understood conveniently.

Kepler's laws of planetary motion read as follows:

1. The first law states that all planets *revolve* around the Sun, the star of our Solar System and they do this on well-defined and stable orbits. These orbits are ellipses as explained and mathematically proven by Kepler's first law published in *Astronomia nova*.
2. The second law says that the speed of the planet on its elliptical orbit changes at each moment such that the time between two positions is always proportional to the area swept out on the orbit between these positions (published in *Astronomia nova*).
3. The third law states that the square of the orbital period P of a planet is proportional to the cube of the semi-major axis a of its orbit

$$P^2 \propto a^3 \tag{1.1}$$

Based on Kepler's third law, it is not surprising to see that with increasing distance from the Sun, revolution periods become smaller. And one implication of the first two laws is that on their respective orbit planets as well as natural and artificial satellites are sometimes closer to their central body and sometimes farther away and their velocity in orbit changes proportionally to the distance of the central body. The point of closes approach is called the *periapsis* Q, while the opposite point is called the *apoapsis*, q.

The sum of periapsis and apoapsis distance $(Q + q)$ equals the major axis, or twice the semimajor axis:

$$Q + q = 2a \tag{1.2}$$

Orthogonal to the ellipse's semimajor axis (a) is the semiminor axis (b). Their length describe the eccentricity e of a planet's orbit

$$e = \frac{q - p}{p + q} \tag{1.3}$$

Kepler defined seven elements to describe a planetary body's motion in space. They can be grouped into elements describing the position and shape of an orbit

(orbit inclination i, eccentricity e), the position of an object on that orbit (ascending node Ω, argument of perigee ω, mean motion v and mean anomaly M), and, the time at which the description of position was valid, the Epoch Time T.

To make calculations more handy, distances in the Solar System are not calculated in kilometres, but in Astronomical Units (AU). The AU was defined as the Earth's semimajor axis, i.e. the average distance between the Sun's centre and Earth's centre, and corresponds to roughly 150 Mio kilometres. In 2012 the *International Astronomical Union (IAU)* has fixed the value to

$$1\text{AU} = 149,597,870,700\,\text{m} \tag{1.4}$$

The so-called *inner planets* are those planetary objects that are closest to Earth, both in terms of position as well as composition and encompass the objects of Mercury, Venus, the Earth and Mars. The Inner Solar System is therefore composed of the inner planets as well as three planetary satellites: the Earth's moon and the Martian moons Phobos and Deimos. Consequently, the *outer planets* in the *Outer Solar System* are the gas planets Jupiter, Saturn, Uranus and Neptune which are accompanied by a large number of moons (see Figs. 1.5 and 1.6).

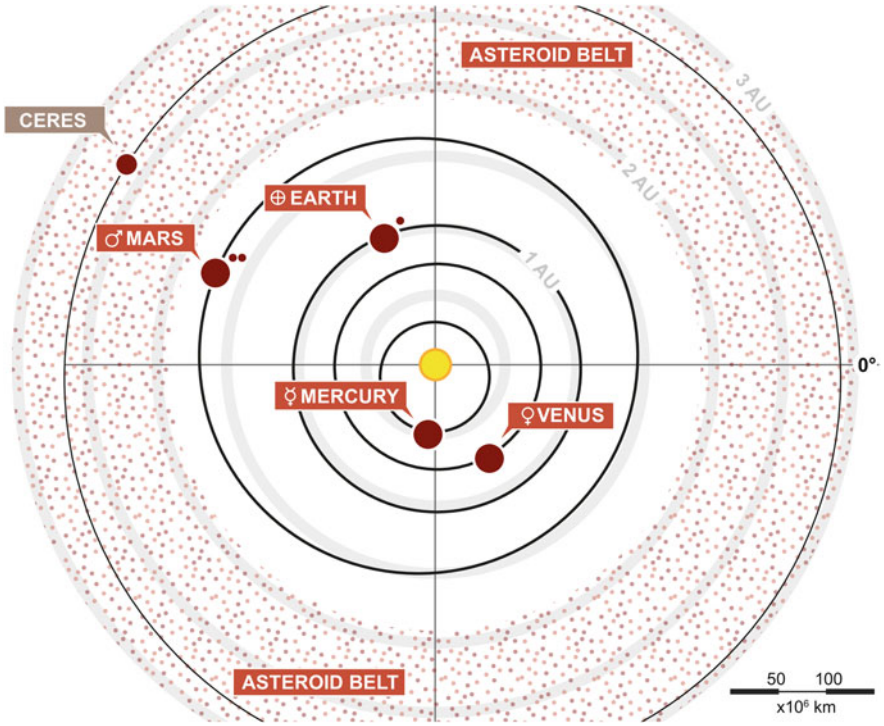

Fig. 1.5 The Inner Solar System with orbits of planets and moons; dwarf planets are colorized in *brown*, solid-surface planets in *red* and gas planets in *blue*

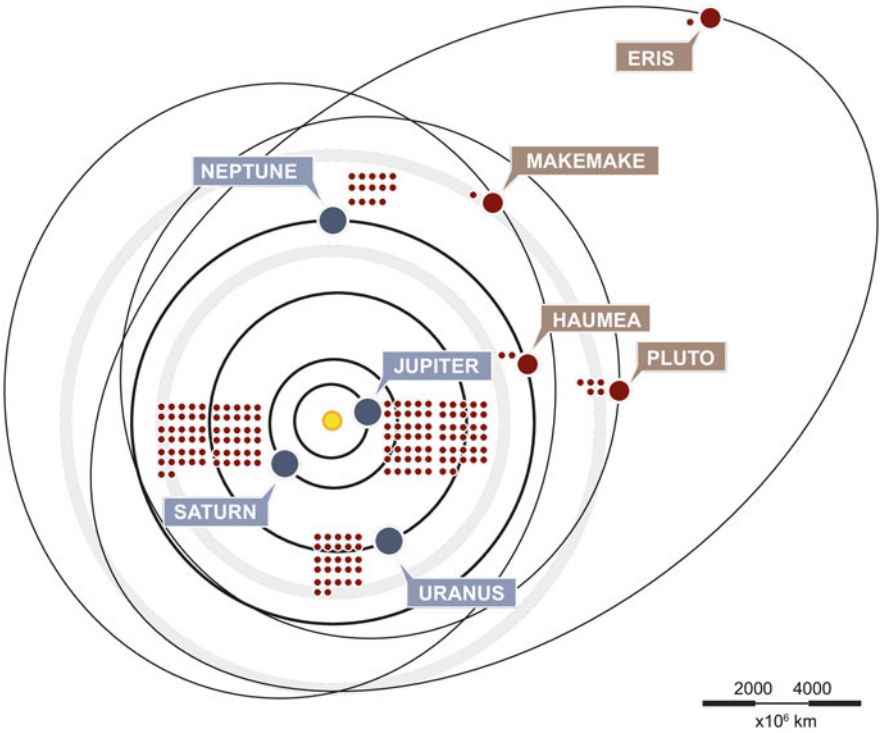

Fig. 1.6 The Outer Solar System with orbits of planets and moons; dwarf planets are colorized in *brown*, solid-surface planets in *red* and gas planets in *blue*

By definition, all objects beyond the orbit of Neptune, i.e. with a>30 AU are called Trans-Neptune objects (TNO) of which the Scattered Disk Objects (SDO) are a subset that is considered to be directly influenced by Neptune's presence (see Fig. 1.7) and a potential source for short-period comets. Also minor planets called *Centaurs*, located along the orbits of Outer planets might belong to the same group as SDOs. Pluto and the other Trans-Neptunian Dwarf Planets Makemake, Haumea and Eris belong to the group of Kuiper-Belt objects (KBO).

Between 1000 and 100,000 AU the Öpik–Oort Cloud extends (see Fig. 1.7) which is considered to be the source area of long-period comets consisting of icy objects, that might enter the Inner Solar System from time to time. It forms the outer boundary of the Solar System and is not considerably influenced by the Sun anymore.

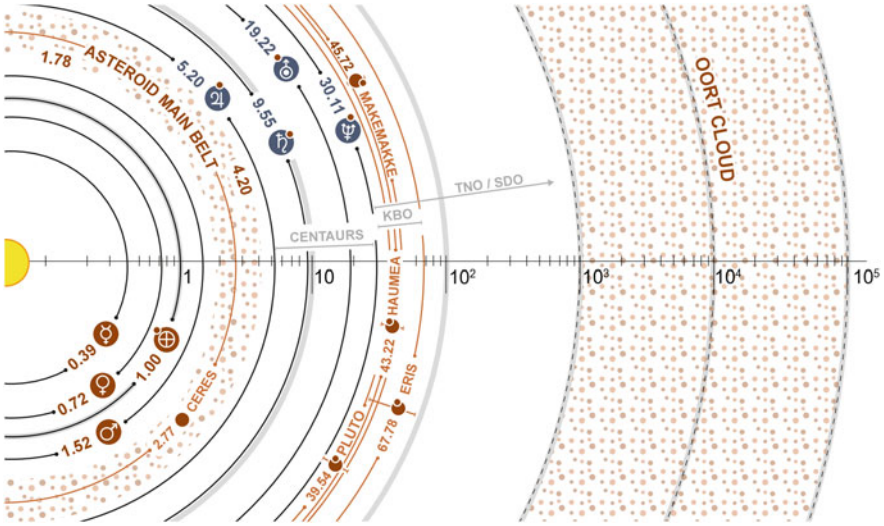

Fig. 1.7 The Solar System between the Sun and the Oort Cloud at 10^5 Astronomical Units. TNO refers to Trans-Neptune objects, SDO refers to Scattered Disk Objects and KBO refers to Kuiper-Belt Objects

1.3 Future Prospects

Predicting the future of discoveries is—by definition—a pointless exercise. What can be expected, however, is the set of missions that are going to be developed withing the next decades. In this respect, Planetary Geology and planetary science in general can be predictable in terms of where we will go, provided that missions will be successfully built, delivered and deployed. What can also be predicted are areas of potential expansion, that in fact drive the requirements for future planetary or space exploration missions.

Mars and the Earth's moon will continue to be prime targets for future in-situ analysis, for the potential establishment of future human bases, for investigating sample return options, for studies on the feasibility of resource extraction and for the investigation of fundamental research related to the geologic evolution and life.

The seemingly increased push towards human exploration, at least of Mars, will have to deal with difficulties of technical nature at all levels (from propulsion, to life support and in-situ resource utilization), but as the Moon exploration with Apollo testifies, human exploration has large advantages in terms of flexibility and short-loop response.

Exobiology and the close investigation of the intricate interplay between geological and biological interaction and the co-evolution of life on the Earth and potentially other bodies in the Solar System continue to be a research topic of top relevance. New instruments will be developed and deployed and beyond Mars future targets will include investigating the solid-surface satellites of Jupiter and Saturn.

What is very different nowadays compared to the 1960s are the developments in terms of robotics, autonomy and, in perspective Artificial Intelligence . All of those, either on their own or matched with human-based geological exploration. In case a faster pace of exploration by humans in the inner Solar System will be imprimed, e.g. in the case of Mars in the forthcoming years, Planetary Geology will certainly have much to gain.

Exoplanetary geology is likely far ahead, but mainly indirect geological evidences on exoplanetary atmospheres, such as on *current* activity at the time when the imaging are delivered to us, (Chaps. 8–10), will probably drive more modelling efforts, before targeted surface imaging will be widely available. Nevertheless, recent space astronomical observatory imaging of rocky exoplanets , in some fortunate case due to both orbital settings and observational geometry, might have allowed to map surface temperatures compatible with partially molten surfaces like that on a large-scale version of Jupiter's satellite Io (Chap. 12). The detection of terrestrial exoplanets, which was impossible only few decades ago, is now ramping up and several missions concur to the discovery of a growing number of candidates. Even the possibility of large numbers of rogue planets, not associated with any central Sun, widens even further the perspectives.

What applies to Earth geology is also valid for Planetary Geology: the present is key to the past (Chap. 2), and our knowledge of Earth and past and present processes are the basis for our interplanetary uniformitarianism, with its assets and its limits.

Further Readings

Carr, M.H.: Geologic exploration of the planets: the first 50 years. Eos **94**(29), 29–30 (2013). doi:10.1002/2013EO030001.

Committee on the Planetary Science Decadal Survey and National Research Council: Vision and Voyages for Planetary Science in the Decade 2013–2022, 400 pp. National Academies Press, Washington (2012) doi:10.17226/13117. www.nap.edu/catalog/13117/

Evans, J.: The History and Practice of Ancient Astronomy, 496 pp. Oxford University Press, Oxford (1998)

Geiss, J., Rossi, A.P.: On the chronology of lunar origin and evolution. Astron. Astrophys. Rev. **21**, 1–54 (2013). doi:10.1007/s00159-013-0068-1

Greeley, R., Batson, R.M. (eds.): Planetary Mapping. Cambridge Planetary Science Series, vol. 6, 310 pp. Cambridge University Press, New York (1990)

Kieffer, H.H., Jakosky, B.M., Snyder, C.W.: The planet Mars: from antiquity to the present. In: Kieffer, H.H., Jakosky, B.M., Snyder, C.W., Matthews, M.S. (eds.) Mars, pp. 1–33. University of Arizona Press, Tucson (1996)

Rossi, A.P., van Gasselt, S.: Geology of Mars after the first 40 years of exploration. Res. Astron. Astrophys. **10**, 621–652 (2010). doi:10.1088/1674-4527/10/7/003.

Sheenan, W.: The Planet Mars: A History of Observation and DisCoVeRy, 270 pp. The University of Arizona Press, Tucson (1996)

Shevchenko, V.V.: Modern Selenography, 288 pp. Nauka Press, Moscow (1980)

Spudis, P., Pieters, C.: Global and regional data about the Moon. In: Heiken, G.H., Vaniman, D.T., French, B.M. (eds.) Lunar Sourcebook, Chap. 10, pp. 595–632. Cambridge University Press, New York (1991)

Tinetti, G.: Galactic planetary science. Philos. Trans. R. Soc. A **372**, 1–15 (2014). doi:10.1098/rsta.2013.0077

Veverka, J.: Planetary Geology in the 1980s. NASA Special Publication, SP-467, xiv+187 pp. NASA, Washington (1983)

Whitaker, E.A.: Selenography in the seventeenth century. The general history of astronomy. In: Taton, R., Wilson, C., Hoskin, M. (eds.) Planetary Astronomy from the Renaissance to the Rise of Astrophysics. Part A: Tycho Brahe to Newton. The General History of Astronomy, Chap. 8, pp. 119–143. Cambridge University Press, Cambridge (1989)

Wihelms, D.E.: To a Rocky Moon: A Geologist's History of Lunar Exploration, xxi–477 pp. University of Arizona Press, Tucson (1993)

Wilhelms, D.E.: The Geologic History of the Moon. U.S. Geological Survey Professional Paper, vol. 1348, viii+302 pp. United States Geological Survey (USGS), Washington (1987)

Chapter 2
Geologic Tools

Monica Pondrelli, Victor R. Baker, and Ernst Hauber

2.1 Geological Reasoning in Planetary Science

Reasoning in geology emerged from the bottom up. Early natural philosophers (the word *scientist* did not exist until it was coined in 1833 by the Cambridge mineralogist, WILLIAM WHEWELL) observed and interpreted the rocks and landscapes that they encountered close up. They learned to interpret these features using synthetic reasoning, that is, the continuous activity of comparing, connecting, and putting together thoughts and perceptions. Their focus was on the formulation of genetic hypotheses that would indicate the causes of the phenomena of interest, and they looked to the consequences of adopting these hypotheses as the means of testing their fruitfulness in a quest for understanding. Much of the needed synthesis was achieved by using geological maps to summarize and communicate the temporal and spatial relationships for rocks and landforms.

The space age has meant that geological reasoning about extraterrestrial planetary bodies must proceed from the top down. Other rocky planetary surfaces are first seen globally and at low resolution. Only later do finer details emerge as spacecraft sensors focus on smaller areas at higher resolution. Eventually landers and robotic rovers image and analyze at the same scales at which geologists originally studied their home planet (Chap. 5). However, unlike the experience with

M. Pondrelli (✉)
Universitá d'Annunzio, Chieti-Pescara, Italy
e-mail: monica@irsps.unich.it

V.R. Baker
University of Arizona, Tucson, AZ, USA
e-mail: baker@arizona.edu

E. Hauber
German Aerospace Center (DLR), Berlin, Germany
e-mail: ernst.hauber@dlr.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_2

studying Earth, this kind of detail only becomes available at a small number of discrete locations, thereby providing less local detail with which to check broader syntheses of geological relationships.

2.1.1 The Problem of Convergence (Equifinality)

The top-down directionality of planetary exploration enhances an explanatory challenge that has been termed *convergence* or *equifinality*, in which the geologist must contend with the possibility that similar effects (landforms, structural patterns, etc.) may be generated by different combinations of causative processes. A classic example of this problem in the history of planetary geology is the debate over the origin of the Moon's crater-like landforms that were first observed in low-resolution telescopic views. Did these features result from explosive volcanism or from meteor impacts? Planetary geology works at resolving such questions though the combination of increased resolution and the study of terrestrial features of known origin that can serve as analogs to the extraterrestrial features of unknown origin. Instead of an equifinality of lunar craters being formed by different kinds of processes, terrestrial analog studies, combined with increased resolution of details, showed that most lunar craters resulted from impact processes. Subsequent work elaborated upon the detailed mechanics of the geologically relevant processes, further strengthening the explanation.

2.1.2 The Role of Analogies

All science relies upon the use of analogy, where *analogy* implies similarity among like features of two otherwise different things. Computational simulations illustrate a form of analogy in which attributes presumed to be fundamental to the two things being compared (attributes such a basic physics or mathematical structure) are incorporated into a simplified system that can then be tested via correspondence to the *real world*. Geological reasoning uses a form of analogy that takes advantage of observed natural regularities that suggest reasoning in which the newly discovered phenomena are compared to phenomena already known and understood. In this way insights gained from the comparison contribute to further investigations into the cause(s) of the unknown phenomenon. Geological analogies serve not so much to provide definitive explanations as they do to provide a source for hypotheses that move geological research into productive lines of inquiry.

2.1.3 Terrestrial Analogs in Planetary Geology

Unlike the newly discovered geological phenomena on other planets, geological phenomena on Earth are much more likely to have both their key features and their causes known. Thus, by recognizing key features shared between terrestrial analogs and their likely extraterrestrial counterparts, geologists can infer likely potential causes for the latter through their understanding of the terrestrial causes.

For experienced geologists this form of inference from analogs is not a trivial *look–alike* exercise. The geologist uses a broad basis of experience with terrestrial phenomena to formulate *multiple working hypotheses* that are evaluated by exploring their consequences relative to a synthesis of understanding.

GROVE K. GILBERT in 1886 pointed out that geologists are investigators rather than theorists. The latter test their theories by assessing the correspondence of theoretical consequences (usually *deductions* or *predictions* from the theories) with specific observations. Ideally such observations are made in the course of controlled experiments. However, in most cases the subject matter of geology is not conducive to the kinds of controlled experiments that characterize much of physics and chemistry. The geologist/investigator cannot place a glacier, a volcano, or an evolving mountain range in a laboratory room. In their scientific reasoning geologists must be more like detectives than laboratory analysts. In the testing of a working geological hypothesis the geologist considers its consequences through their *consistency*, *coherence* and *consilience* with related phenomena.

Consistency entails a lack of contradiction, such that a causative hypothesis for a geological phenomenon is not contradicted by an indicated historical sequence of development or spatial relationships with other phenomena. *Coherence* requires an explanation that is sufficiently comprehensive to align with other known explanations of closely related phenomena. Finally, a tentative hypothesis can be evaluated in terms of *consilience*, literally, a *jumping together* of knowledge, if it leads to a kind of *explanatory surprise* in which a completely different set of phenomena from that being tentatively explained is discovered or recognized, such that (1) the newly recognized phenomena are clearly related to the phenomena under investigation, and (2) that they are adequately explained by the tentative hypothesis that was originally proposed in a more limited context. Though consilience does not confer truth via formal logic, its operation has long been recognized to be associated with the most fruitful of scientific investigations.

While analogies do not by themselves provide complete explanations for planetary landform genesis, they may initiate a line of inquiry that places the investigator on a reliable path toward those explanations. The investigator presumes a well-reasoned analogy as possibly true, and then infers what would be expected consistent with that presumption. In practice, the classification of phenomena, the recognition of potential analogues, and the corroboration of working hypotheses via consistency, coherence, and consilience all occur during the course of regional planetary mapping. The mapping process itself allows the geologist to continually assess hypotheses for a feature's cause and age by checking these against the mapped relationships.

2.1.3.1 Examples of Terrestrial Analogues

Classical examples of early terrestrial analogues were impact craters and basaltic volcanic provinces. For example, the Apollo astronauts were trained in the Ries impact crater in Germany to prepare for investigations of lunar impact structures and related rocks (Fig. 2.1). First–hand experience in the recognition of volcanic rocks and their petrography were obtained by the astronauts in locations such as Iceland, where newly erupted basaltic lava flows and related deposits are abundant. Impact craters on other planetary surfaces could be observed early with telescopes and impact craters on the Earth were among the first and most intensely studied analogues. Among the most frequently studied impact analogues were the Ries crater in Germany, Meteor Crater in Arizona, U.S., Lonar Crater in India, and the Haughton Crater on Devon Island, Canada.

Analogue studies had their next peak period after it had become clear through the analysis of data returned by the Mariner 9 mission, that Mars was not a dead and dry place like the Moon. Instead, the Mariner 9 images showed a plethora of landforms that bear evidence of endogenic and exogenic activity, such as volcanic, fluvial and

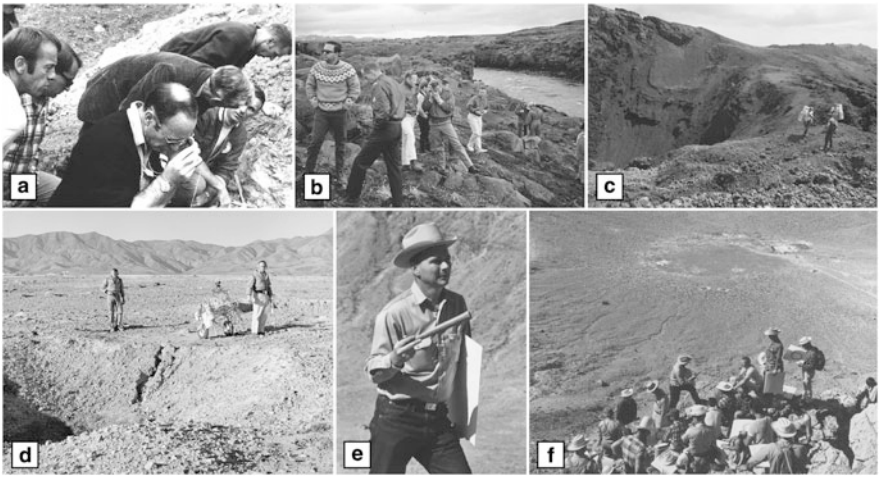

Fig. 2.1 Geological field training for Apollo astronauts was conducted at various terrestrial analogs. (**a**) Quarry at Otting (Nördlinger Ries, Germany); the Ries is an easily accessible, large impact crater that was a convenient analog for lunar craters; *from the left*: A. SHEPARD, F. HÖRZ, E. MITCHELL, W. VON ENGELHARDT, G. CERNAN, and J. ENGLE; (**b**) astronauts on a field excursion on Iceland. Iceland offers easy access to basaltic volcanic landscapes and was considered by some as the most lunar-like place during Apollo crew training; (**c**) the Apollo 15 crew conducts geological training in Apollo Valley on Hawaii's Big Island; (**d**) astronauts A. SHEPARD and E. MITCHELL prepare for Apollo 14 at an artificial crater field in Arizona; (**e**) G. SHOEMAKER, one of the pioneers of planetary geology, was instrumental in Apollo crew field training; (**f**) G. SHOEMAKER (with hammer) lectures to astronauts at Meteor Crater, Arizona, another frequently used terrestrial analog to planetary impact craters (crater floor on top). Source: (**a**) D. Stöffler/NASA. (**b**)–(**c**) NASA. (**d**)–(**f**) USGS

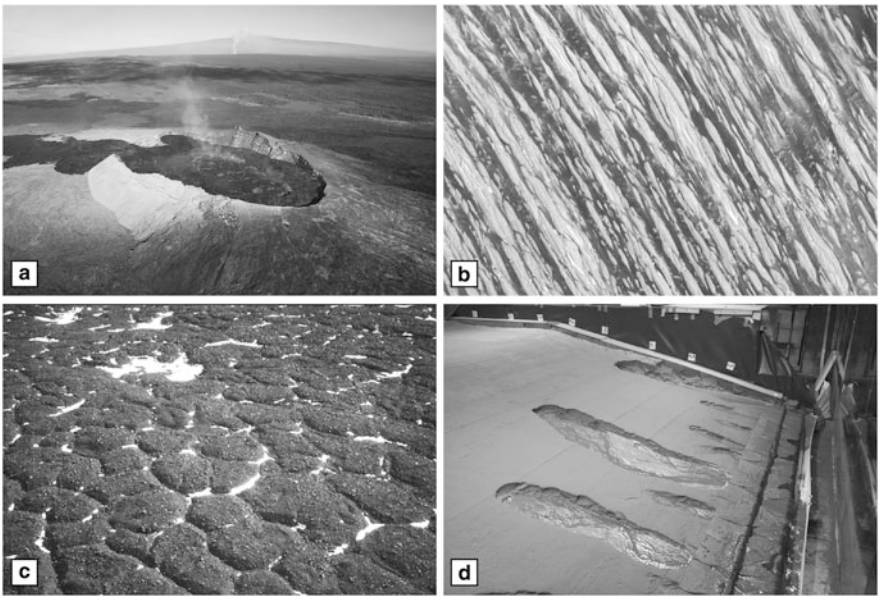

Fig. 2.2 Examples of Earth analogues for Mars research; (**a**) the Pu'u 'Ō ō vent on the slopes of Kilauea, Hawaii, with the huge shield volcano, Mauna Loa, in the background; (**b**) yardangs in the Dasht-e Lut desert (Iran); (**c**) patterned ground (sublimation polygons) in Beacon Valley, part of the McMurdo Dry Valleys in Antarctica; (**d**) groundwater seepage experiments in the *Total Environmental Simulator* facility of the University of Hull (UK). Source: (**a**) USGS. (**b**) NASA. (**c**) D. Marchant/NSF. (**d**) W. Marra/University of Utrecht

eolian processes. These discoveries led to the development of the Viking mission, the main goal of which was to search for life on the surface of Mars. In preparation of Viking, NASA funded several comprehensive comparative planetology field guides and reports that focused on the remote sensing analysis of landforms that are analogous to features that were identified in Mariner 9 images. Hence, these analogues were based on morphological similarity. Basaltic volcanic lava flows and shield volcanoes in Hawaii (Fig. 2.2a) and the Snake River Plains, Idaho, U.S., served as analogues for volcanic flows and edifices in Tharsis, the largest volcanic province on Mars.

Landforms typical for arid climates were studied in the Egyptian desert, and a class of streamlined erosional landforms on Mars that display a characteristic *inverted boat hull* shape were compared to yardangs in Peru (Fig. 2.2b). The discovery on Mars of deeply incised channels that have an abrupt, amphitheater-shaped heads and only few tributaries sparked the question which aqueous process was responsible for this peculiar morphology. Analogous valleys on the Colorado Plateau in the southwestern United States were, at that time, thought to have formed by sapping, that is, backward erosion by seepage (the emergence of groundwater at the foot of a cliff). Perhaps one of the most enlightening applications of analogous

reasoning in planetary geomorphology led to the hypothesis that enormous fluvial channels on Mars may have formed by catastrophic floods. Morphologically similar, yet smaller channel systems in Washington State, U.S. were formed by several such catastrophic floods, which were triggered by the sudden drainage from glacially dammed lakes in the Pleistocene (Missoula floods). Iceland hosts volcanic and glacial landscapes and is one of the best sites to study volcano-ice interactions, which are thought to have been important on Mars, too.

The analysis of radar observations showed that Venus also has a surface that displays a rich inventory of landforms. Volcanic and tectonic features were most abundant, but wind-related phenomena were also ubiquitous. Analogues for dunes, yardangs and wind streak-patterns were investigated in the United States and other places such as Chad and Bolivia. To maximize the return of these analogue studies, they were mainly based on radar observations of terrestrial features, as the set of parameters influencing them (e.g., viewing direction, incidence angle, and polarization, but also factors such as moisture content) are different from those parameters that are important in the interpretation of visible images. Venusian volcanic landforms were compared to a variety of terrestrial analogues. For example, small volcanic domes were compared to basaltic lava shields on Iceland, on the basis of their shapes and volumes. Of particular interest for the emplacement of volcanic landforms is the unique surface environment of Venus, which is characterized by very high temperature and atmospheric pressure. The terrestrial volcanic environment with the highest ambient pressures is the seafloor; hence a number of analogue studies used seamounts and other submarine volcanic features as analogues to Venusian volcanic landscapes.

The next peak in research on terrestrial analogues was triggered by the trend to increasing image resolution of camera systems. Whereas the Viking Orbiter images of the 1970s have typical pixel sizes of 60 m and could resolve landforms that are mostly kilometer-sized or larger, more recent imaging systems such as the ESA Mars Express High Resolution Stereo Camera (HRSC), Mars Reconnaissance Orbiter Context Camera (CTX), and High Resolution Imaging Science Experiment (HiRISE) have ground pixel sizes of $\sim$10 m, 6 m, and $\sim$30 cm, respectively. The study of these high-resolution images led to new discoveries, e.g., the geologically young gully systems on Mars, which are characterized by much larger spatial scales than the volcanoes or outflow channels of the Viking era. The new or refined identifications of, e.g., periglacial surface features, required additional terrestrial analogues that were not considered during the first, Mariner 9 and Viking-phase of Mars exploration (Chap. 9). Gullies and alluvial fans were studied in diverse locations on Earth, including the Atacama Desert or Svalbard, Norway. When it became obvious that the mid-latitudes of Mars are characterized by a diverse set of possibly ice-related landforms that are hypothesized to be the result of climate fluctuations, cold climate region on Earth provided useful analogues. The Tuktoyaktuk peninsula in northwest Canada is a prime site to study pingos, which have been hypothesized to exist in some places in the northern hemisphere and

the Argyre region of Mars. Patterned ground has been identified on Mars at many locations, and was compared to analogue sites in permafrost areas in northern Canada and elsewhere in the Arctic. Representing the most extremely cold, dry land–surface environment on Earth, Antarctica has been selected as a terrestrial analogue for Mars since the early 1970s. Specifically, the hyperarid polar desert landscape of the McMurdo Dry Valleys (Fig. 2.2c), which are also relatively easy to access, was very instructive in the interpretation of recent and ancient Martian environments.

With increasingly sophisticated payloads, more and more compositional information of planetary surfaces becomes available. Therefore, analogue studies are also becoming more diverse, and continue to include an ever increasing variety of geochemical and mineralogical examples of terrestrial environments. Most of them have been chosen for Mars research, as orbiting spectrometers and, increasingly, in situ investigations by landers and rovers provide unprecedented details on the composition of surface materials. For instance, a lot was learned about the possible formation conditions of the iron oxides and sulphate minerals found by the MER rover Opportunity, in Meridiani Planum, from the unique acidic alteration environment at Rio Tinto in southern Spain. After the discovery of perchlorates in Martian soil by the Phoenix lander, scientists turned to the Atacama Desert in Chile, as it is there where the highest perchlorate concentrations on Earth are measured. Another frequently used geochemical/mineralogical analogue is provided by the hydrothermal environment of the Yellowstone caldera, U.S., where silica-rich soils can be studied (Chap. 14).

Analogue studies are not limited to natural examples. Recently, more and more studies use laboratory experiments to investigate processes assumed to operate on planetary surfaces. Some of them can be reasonably well simulated to evaluate which are the relevant parameters that control their formation. One example is sapping, i.e. the backward erosion triggered by groundwater seepage. This process was long thought to be possibly responsible for the origin of a particular type of Martian valleys, but recent terrestrial studies have shown that groundwater alone does perhaps not provide the erosive capacity to incise the respective valleys, examples of which are abundant on the Colorado Plateau in Utah and Arizona, U.S. Scaled landscape development experiments can help to determine which characteristic landform elements are diagnostic of specific processes (Fig. 2.2d). Certain processes are sensitive to specific environmental parameters such as gravity and pressure. The lower gravity on Mars cannot be permanently simulated on Earth, but experiments in low gravity test facilities such as parabolic flights or drop towers can help to quantify gravity effects on geologic processes and relevant parameters (e.g., the angle of repose). The low temperatures and pressures on Mars are beyond any natural conditions on Earth. Mars Simulation chambers can, to some degree, help to create such conditions. For example, it was recently tried to reproduce flow phenomena leading to the formation of debris flows and recurrent slope lineae (RSL).

Most terrestrial analogues have been applied to the study of the terrestrial planets. The surfaces of the icy satellites in the Outer Solar System are unlike anything on Earth, hence it is not easy to find analogues. Nevertheless, some basic aspects of their surfaces and interiors can be compared to natural phenomena on Earth. For example, the subglacial lakes in Antarctica (e.g., Lake Vostok) may be useful analogues to the sub-ice oceans on Europa or Enceladus, and it will be interesting to see whether they host unique habitats, as it is hypothesized by some researchers for the icy satellites as well. Other properties of the crusts of the icy satellites can only be investigated in the laboratory. At very low temperatures, ice does behave mechanically quite different than on Earth (almost like rock), and this needs to be studied experimentally. It is beyond the scope of this chapter to provide a comprehensive list of all types of terrestrial analogues used so far. With the ongoing pace of planetary exploration, analogue studies remain as important as ever. The increasing importance of landed missions and in situ investigations will require that more analogues are identified that help to understand compositional observations at small scales and specific geochemical environments.

2.1.4 The Stages of Geological Reasoning

Analogy is essential to geological reasoning. As recognized by GILBERT in the late nineteenth century, broad experience with and understanding of terrestrial geological phenomena provide geologists with their most effective resource for the invention of potentially fruitful, working hypotheses. The actions of (1) forming such hypotheses, (2) following their consequences, and (3) testing those consequences comprise integral parts of effective geological reasoning in regard to the understanding of planetary surfaces. Following the formulation of hypothesis or hypotheses (some of which may regenerate from or replace the initial ones), inquiry can then be further advanced through the quantitative modeling of various system components.

2.2 Stratigraphy: The Tool to Order Rocks and Time

The word *stratigraphy* derives from the Latin *stratum* and the Greek *graphia*, meaning the description of the rock bodies and their organization into distinctive units based on some of their properties. The aim is to recognize their relative and vertical distributions and relations in order to infer their succession in time (i.e., the sequence of events) and interpret the geological history of a given area and, in perspective, of a planet. As a consequence, stratigraphy is a fundamental and basic science for any geological reconstruction, at all spatial and temporal scales.

Stratigraphy has its roots in the Renaissance period, starting with the rediscovery of work of Greeks and Arabs in natural philosophy. In particular the rediscovery of

geometry was fundamental to grasp the basic geological principles since geology is based on spatial relations. LEONARDO DA VINCI (1452–1519) understood some of the basic principles of sedimentary process and stratigraphy. GEORGIUS AGRICOLA (1494–1555) illustrated layers and observed how they could be traced over a wide area. NICOLAUS STENO (1638–1686) was the first in 1669 (at least the first whose record is left) to organize, write and apply to a rock outcrop the principles, which ever since have been known as Steno's principles of stratigraphy.

The *principle of superposition* states that in an undisturbed succession of layers the oldest rocks are at the base of the succession and the youngest are at the top. As an example, in Fig. 2.3 the effusive igneous rocks (A) represent the oldest unit, followed by the evaporite layers (B–G) and then the sandstones (H).

The *principle of original horizontality* states that layers form in horizontal position and only later are eventually brought in other positions by tectonic forces or, such as in the case pictured in Fig. 2.4, gravitative-driven processes (C). Exceptions to this rule occur in particular environments where layers form following the angle of repose of the material (clinoform, see Sect. 2.2.2).

The *principle of lateral continuity* states that each layer would be continuous throughout the Earth surface unless it meets some solid body (boundary of the depositional basin). In Fig. 2.4, the impact breccia (B) and the evaporites (C) terminate against the crater margin (A) which represents the lateral limit of their

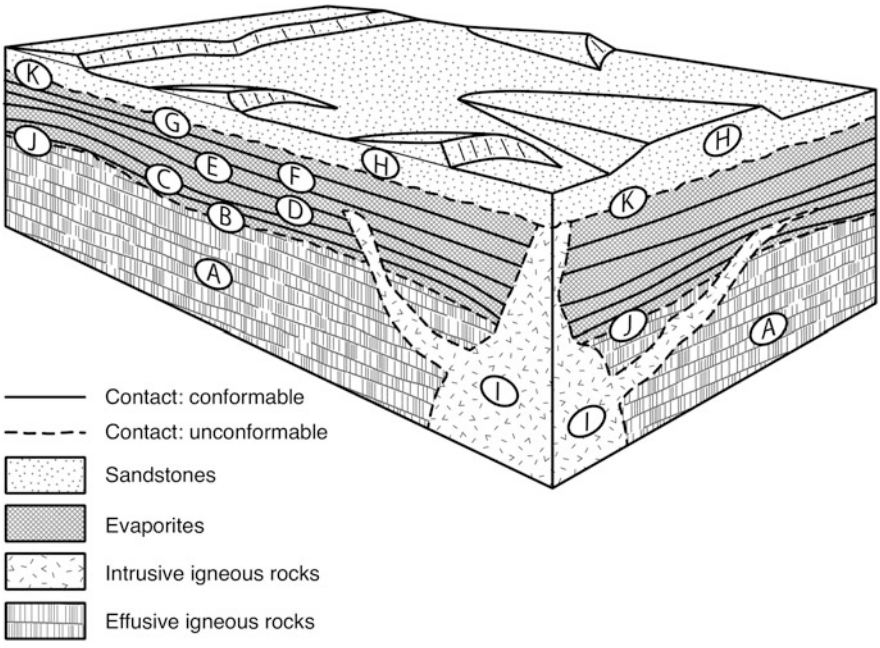

Fig. 2.3 Block diagram showing some stratigraphic relations between layered and non-layered rocks and exemplifying the principles of stratigraphy. See text for explanation

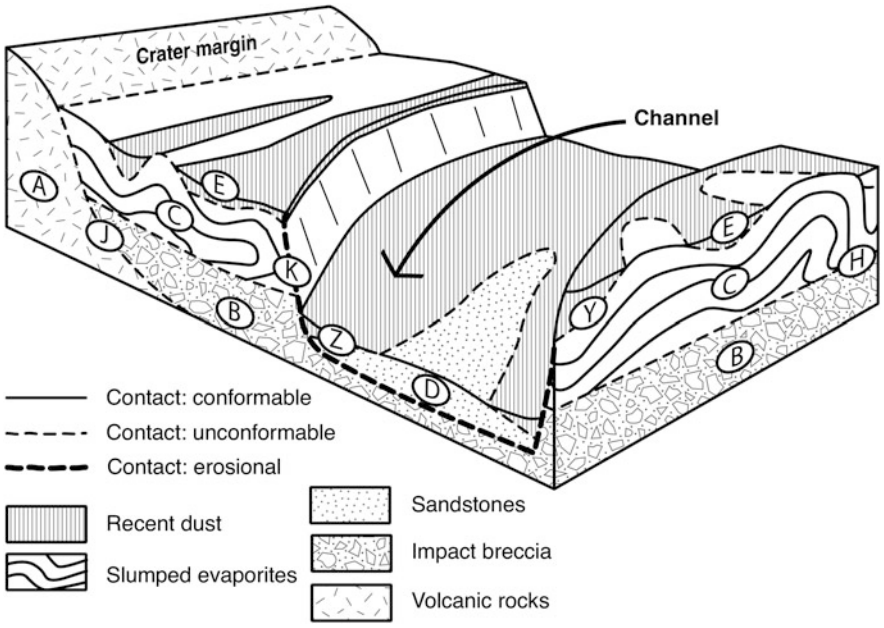

Fig. 2.4 Block diagram showing some stratigraphic relations between layered and non-layered rocks and exemplifying the principles of stratigraphy. Unit C consists of layers of evaporites which have been folded and deformed by a slump, losing their original horizontality

deposition. Instead, the same units B and C are cut by the channel that interrupts their original lateral continuity.

The *principle of cross-cutting* relationships states that if a layer is cut by another rock or a discontinuity, the latter must have formed after the layer. In Fig. 2.3, the intrusive igneous rocks I cut and thus postdate the effusive igneous rocks (A) and the evaporite beds (B–G). On the other hand, the sandstone unit H is not affected and cover, thus cuts and postdates the unit I. In Fig. 2.4 the channel cuts and postdates the impact breccia (B) and the evaporites (C) while the recent dust (E) covers the other units both outside and inside the channel thus postdating it.

2.2.1 Relative Stratigraphy

Steno's principles of stratigraphy constitute the basic tools to order the rocks of a specific area with respect to time, reconstructing a relative dating of the geological units. This applies to any planetary surface. This approach allows to identify the relative order in which rocks were deposited and, since each layer or rock represents a piece of history in term of the environment in which was formed, reconstructing a succession of geological units means to reconstruct a succession of events. In Fig. 2.3, applying the principles of superposition and cross-cutting

relationships, the relative stratigraphy can be accordingly reconstructed from older to younger: effusive igneous rocks (A), evaporites (B–G), intrusive igneous rocks (I) and sandstones with dunes (H). In Fig. 2.4, the succession consists of the volcanic rocks forming the crater margin (A) followed by the impact breccia (B) and the slumped evaporites (C); then the sandstones forming the channel deposits (D) and the recent dust deposits (E) which are deposited inside and outside the channels.

An example of stratigraphic reconstruction comes from the Taurus-Littrow region of the Moon, in correspondence of the Apollo 17 landing site. This location is particularly interesting because the remote imagery is coupled with the ground truth performed by the astronauts during the mission (Fig. 2.5a), as well as with the analyses of samples returned to Earth. Here an example of stratigraphic relations

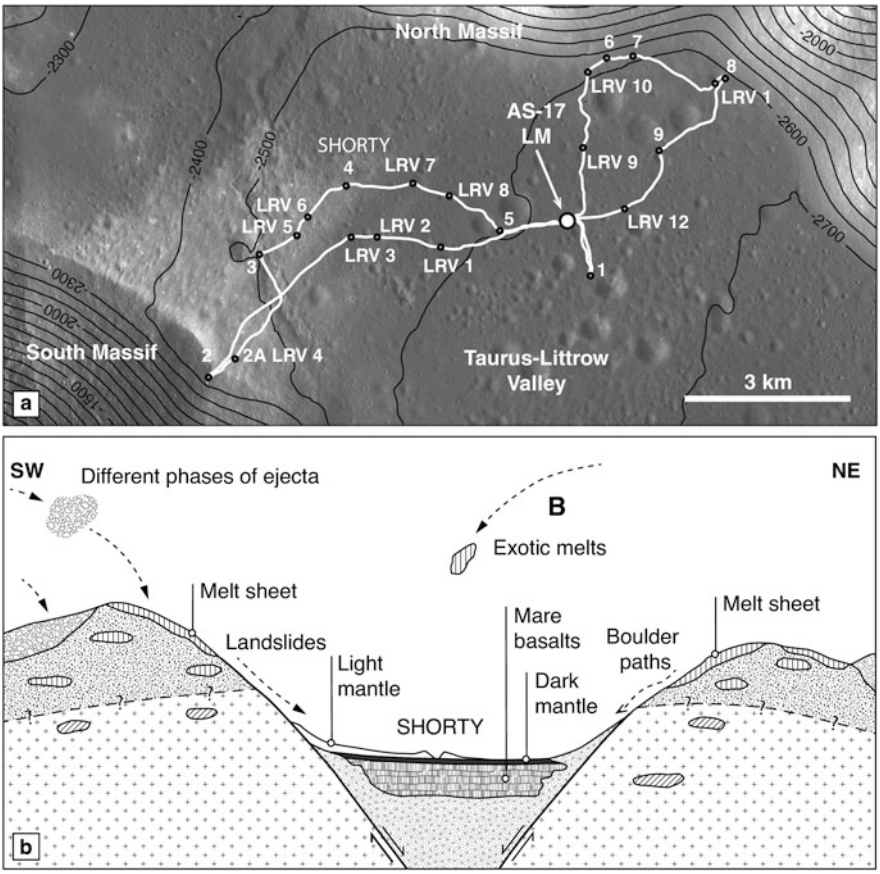

Fig. 2.5 Example of relative stratigraphy reconstruction in the Apollo 17 landing site, Moon; (**a**) NASA LRO LROC image mosaic showing the analysed area; (**b**) Geological section across the landing site area. Source: redrawn after Spudis and Pieters (1991), NASA LRO/NAC image mosaic from M104311715LE, M104311715RE, M180966380LE, M104318871RE, M1142241002RE, M180966380LE

between older highland material and younger mare basalts is present. The massifs represent the highland material and consist of different layers of ejecta and impact–related deposits. These units are embayed by the younger mare basalts, associated with debris material coming from the massifs, that were emplaced in the fault-bounded valley. These relations are sketched in the geological cross section of Fig. 2.5b. The section emphasizes the presence of the extensional faults, probably formed after an impact event, that create the accommodation space for the younger basalts to be deposited. Geologic cross sections represent an effective tool to understand and visualize the vertical and lateral stratigraphic relations between the different units and features (Chap. 4), including faults, in order to infer the geometry of the different depositional bodies.

Lateral transitions illustrate the passage between different morphologies or deposits, which in turn reflects the transition between different environments. This allows the geologist to integrate the single landform or deposit in a framework of more landforms/deposits (Chap. 9), in order to strengthen their genetic interpretation.

An example of lateral transition is given in Fig. 2.6, taken from the Firsoff crater on Mars. Here mounds up to few hundred meters large at their base are associated with sulphate-bearing layered deposits Fig. 2.6a. The perspective view of Fig. 2.6b allows the interpreter to follow the layers from the mounds to the layered deposits adjacent to the mounds. This lateral continuity of the layers implies that they were formed in two different, but adjacent and coeval, depositional environments.

2.2.2 Layer Terminations and Geometries

In order to reconstruct the relative stratigraphy of a given area, the termination of layers and in general of rock units must be carefully evaluated. These geometries were understood on Earth on seismic profiles, but then became of common use in field and remote geology on Earth. The higher resolution data now available for some planetary bodies enable some of these observations on planetary studies as well.

The presence of an erosional truncation implies the deposition of a unit, then a subsequent erosion of part of this unit which generates an unconformity surface, which is a surface that corresponds to a geological interval in which time is not represented within the rock record (hiatus). In Fig. 2.7 an example from the Holden crater on Mars is depicted. The light-toned layered deposits (LLD) are locally eroded and then covered by the dark–toned deposits (Dd).

When horizontal or gently inclined layers terminate against an inclined surface they define an *onlap* geometry, which reflects a progressive infill of the basin. In Fig. 2.8 eroded fluvial deposits (R-1) onlap against more inclined surfaces made of dissected crater ejecta (DiCE) close to Neves crater (Mars).

Deposits with naturally inclined beds, called clinoforms, can form in certain sedimentary systems, such as (although not exclusively) the deltaic system. Those are depicted in Fig. 2.9, located inside Holden crater (Mars).

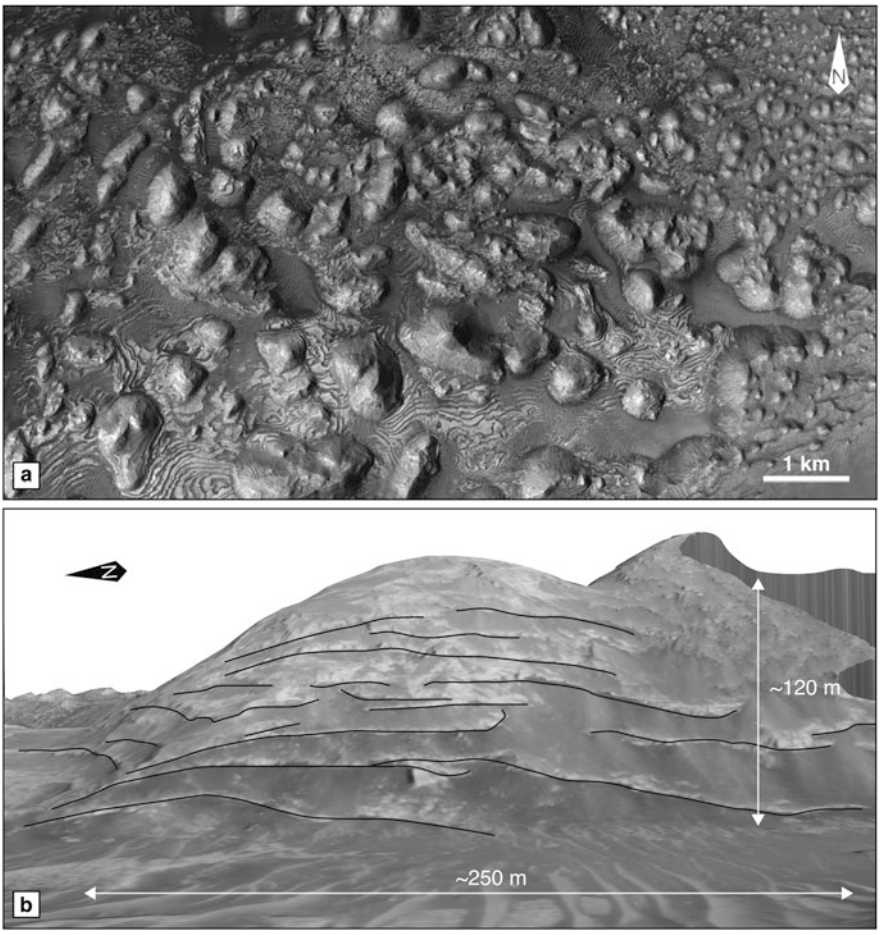

Fig. 2.6 Example of lateral transition between different units from Firsoff crater (Mars); (**a**) the area is characterized by the presence of mounds and layered deposits; (**b**) perspective view of a mound passing laterally to layered deposits. The layers within the mound continue laterally to the layers forming the layered deposits, implying that the two units are coeval. Source: (**a**) NASA MRO/CTX image mosaic, after Pondrelli et al. (2015). (**b**) HiRISE–based DTM (Digital Terrain Model) from images PSP 003788_1820 and ESP 020679_1820

Such a geometry implies that the quantity of sedimentary materials exceeded the available space for deposition, forcing the deposition to occur progressively basin–ward. This pattern is called progradational and, when inclined layers terminate down dip against a sub–horizontal surface, the resulting geometry is called downlap. In Fig. 2.9 the layers of the dark–toned deposits consist of prograding clinoforms which terminate downdip on the light-toned horizontal layers with a downlap geometry.

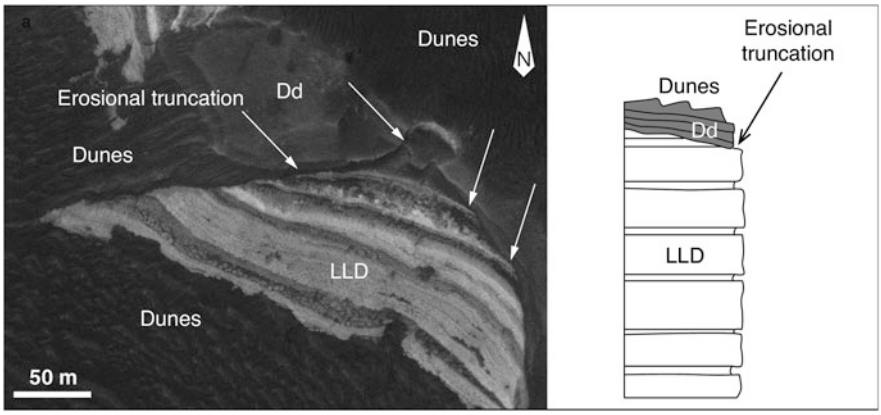

Fig. 2.7 Erosional truncation from the Holden crater (Mars) and sketch. Light-toned deposits (LLD) are truncated as shown by the *white arrows* and then covered by the dark-toned deposits (Ds). This contact implies that between the deposition of LLD and Dd erosion and non-deposition occurred, which in turn implies that the corresponding time is not registered in the rock record. Source: NASA MRO/HiRISE ESP 012386_1530 from Pondrelli et al. (2005), Grant et al. (2008)

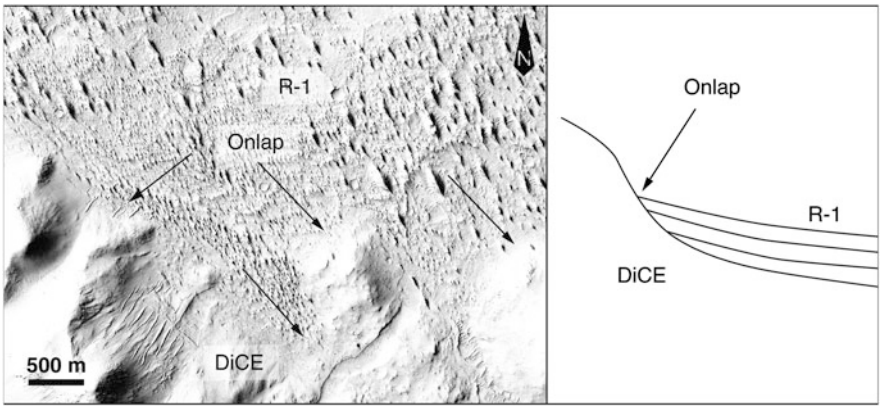

Fig. 2.8 Onlap of eroded fluvial deposits against dissected crater ejecta from the Neves crater (Mars) and sketch. This contact implies progressive infilling of the basin by the R-1 unit. Source: NASA MRO/HiRISE ESP 017047_1770 from Kite et al. (2015)

2.2.3 Unconformities and the Missing Time

A succession of sedimentary layers that have been deposited undisturbed in continuity, originates a conformable sequence.

The erosional truncation pictured in Fig. 2.7 represents an unconformity: An unconformity is an erosional or non-depositional surface separating two rock masses or strata of different ages, indicating that sediment deposition was not continuous. An unconformity implies lack of continuity in the deposition, because

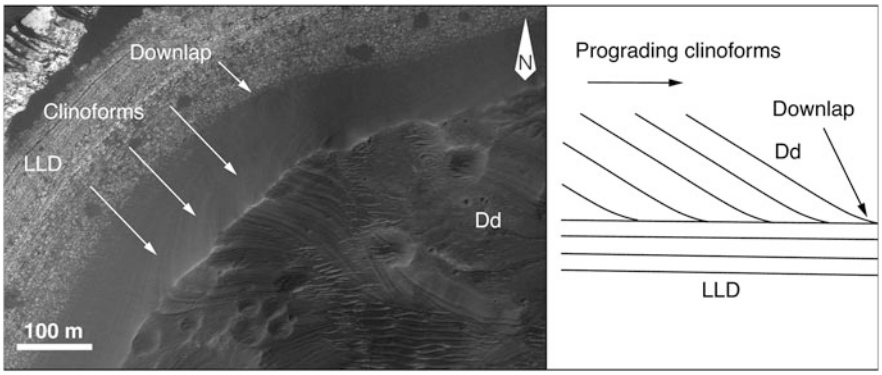

Fig. 2.9 Prograding clinoforms and downlap from the Holden crater (Mars) and sketch. The Dd progrades on top of the LLD, filling the available space for the deposition and progressively depositing basin–ward. Source: NASA MRO/HiRISE PSP 003077_1530 from Grant et al. (2008)

of non–deposition and/or erosion of previously formed layers, before the deposition of younger layers.

An unconformity separating crystalline rocks that have been subjected to erosion and layered sedimentary rocks is termed nonconformity, such as surface J in Fig. 2.3. An unconformity separating layers that are basically parallel and involving significant erosion in addiction to non-deposition, is termed disconformity, such as surfaces K in Fig. 2.3, Z in Fig. 2.4 as well as the erosional truncation of Fig. 2.7. If the unconformity separates rocks whose bedding is not parallel, it is called angular unconformity, like the surface Y of Fig. 2.4. In this case, the unconformity marks an episode of non-deposition, erosion and tilt. A case with non-deposition and not discernible erosion between parallel beds is termed paraconformity, although this feature is so subtle that its recognition with currently available tools in planetary geology is impossible.

The development of a space-time diagram constitutes a tool to express the geological history of a given area through time, emphasizing not the depositional geometry of the different units but rather their duration through time, and, even most importantly, the time which is not represented in the geological record (Fig. 2.10a, b).

Figure 2.10a represents the space-time diagram referred to the block diagram of Fig. 2.3. The missing time, emphasized by the unconfomities, is expressed by the vertical lines. The nonconformity corresponding to the surface J in Fig. 2.3 embodies an interval of time during which no deposition occurred and part of the previously deposited material of unit A was subsequently removed by erosion. The disconformity corresponding to the surface K in Fig. 2.3 represents a time interval of erosion and non deposition of the evaporite layers. The intrusive body I was emplaced and then subjected to erosion during this period. Space time diagrams can be useful to visualize lateral correlation when superposed depositional and erosional events are present (Figs. 2.7, 2.8, and 2.9). For example, the slumped layers C of

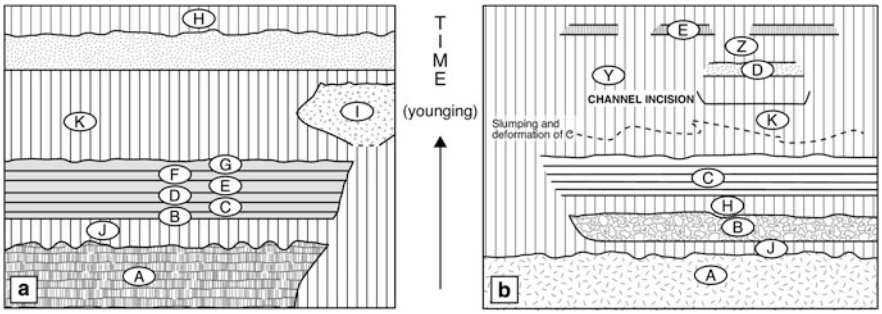

Fig. 2.10 (**a**) Space-time diagram relative to the block diagram of Fig. 2.3 (explanation of symbols there). Units are represented showing their mutual vertical and lateral geometric position compared to their duration though time. The missing time is expressed by the *vertical lines*; (**b**) Space-time diagram relative to the block diagram of Fig. 2.4 (explanation of symbols there). Units are represented showing their mutual vertical and lateral geometric position compared to their duration through time. The missing time is expressed by the *vertical lines*

Fig. 2.3 were laterally extensive before they were cut and partly eroded by the fluvial channel that later was filled by the fluvial deposits D. Then aeolian dust deposits E draped in patches over all the older materials.

Take-Home Messages

Formulation of genetic hypotheses aids understanding the causes of the phenomena of interest.

Consequences of adopting genetic hypotheses are tested through their consistency, coherence and consilience with related phenomena.

Synthesis is achieved by using geological maps to summarize and communicate the temporal and spatial relationships for rocks and landforms.

Analogy is essential to geological reasoning because geological phenomena on Earth are much more likely to have both their key features and their causes known. Recognizing key features shared between terrestrial analogs and their likely extraterrestrial counterparts is essential to infer potential origin and causes for the latter.

Steno's principles are the tools to order the rocks of a specific area with respect to time, reconstructing a relative dating of the geological units: principles of superposition, original horizontality and cross-cutting.

Geological sections provide a tool to visualize the recognized vertical and lateral stratigraphic relations.

Lateral transition illustrates the passage between different but coeval morphologies or deposits which reflects the transition between different environments.

Layer terminations are important for understanding the geometry and evolution of the deposits.

Unconformities are important for understanding the significance of the missing time.

Space-Time diagrams are the tool to represent the temporal and lateral distribution of the different geologic units, including the missing time.

Suggested Readings

Baker, V.R.: Terrestrial analogs, planetary geology, and the nature of geological reasoning. Planet. Space Sci. **95**, 5–10 (2014). doi:10.1016/j.pss.2012.10.008

Beysen, D., van Loon, J.: Generation and Applications of Extra-Terrestrial Environments on Earth. River Publishers Series in Standardisation, 318 pp. River, Aalborg (2015)

Chamberlin, T.: The method of multiple working hypotheses. Science **15**(366), 92–96 (1890). doi:10.1126/science.ns-15.366.92

Garry, W., Bleacher, J. (eds.): Analogs for Planetary Exploration. No. 483 in GSA Special Paper, 567 pp. The Geological Society of America, Boulder (2011)

Gilbert, G.: The inculcation of scientific method by example, with an illustration drawn from the Quaternary geology of Utah. Am. J. Sci. **31**(184), 284–299 (1886). doi:10.2475/ajs.s3-31.184.284

Grant, J., Irwin, R., Grotzinger, J., Milliken, R., Tornabene, L., McEwen, A., Weitz, C., Squyres, S., Glotch, T., Thomson, B.: HiRISE imaging of impact megabreccia and sub-meter aqueous strata in Holden Crater, Mars. Geology **36**(3), 195–198 (2008). doi:10.1130/G24340A.1

Kite, E., Howard, A., Lucas, A., Armstrong, J., Aharonson, O., Lamb, M.: Stratigraphy of Aeolis Dorsa, Mars: stratigraphic context of the great river deposits. Icarus **253**, 223–242 (2015). doi:10.1016/j.icarus.2015.03.007

Koutsoukos, E.: Stratigraphy: Evolution of a Concept. Topics in Geobiology, vol. 23, Chap. 1, pp. 3–19. Springer, Dordrecht (2005). doi:10.1007/1-4020-2763-X_1

Pondrelli, M., Baliva, A., Di Lorenzo, S., Marinangeli, L., Rossi, A.: Complex evolution of paleolacustrine systems on Mars: an example from the Holden crater. J. Geophys. Res. Planets **110**(E4) (2005). doi:10.1029/2004JE002335

Pondrelli, M., Rossi, A., Le Deit, F., Fueten, L., van Gasselt, S., Glamoclija, M., Cavalazzi, B., Hauber, E., Franchi, F., Pozzobon, R.: Equatorial layered deposits in Arabia Terra, Mars: facies and process variability. Geol. Soc. Am. Bull. **127**(7–8), 1064–1089 (2015). doi:10.1130/B31225.1

Spudis, P., Pieters, C.: Global and Regional Data About the Moon. In: Heiken, G.H., Vaniman, D.T., French, B.M., et al. (eds.) Lunar Sourcebook, Chap. 10, pp. 595–632. Cambridge University Press, Cambridge (1991)

Whewell, W.: The Philosophy of the Inductive Sciences: Founded upon Their History, vol. 1, cxx+523 pp. John W. Parker, London (1840)

Winter, J.G.: The Prodromus of Nicolaus Steno's Dissertation Concerning a Solid Body Enclosed by Process of Nature Within a Solid; An English Version with an Introduction and Explanatory Notes, 160 pp. The Macmillan Company, New York (1916)

Chapter 3
Exploration Tools

Stephan van Gasselt, Angelo Pio Rossi, Damien Loizeau, and Mario d'Amore

3.1 Introduction

The geological exploration of the Solar System beyond Earth started with telescopes, and the very first geological observations of extraterrestrial solid bodies were those performed on the Moon. Even the very first published lunar geological maps were obtained from ground-based astronomical observations and were crafted in the early 1960s.

As of today, planetary exploration has been robotic, or jointly human-robotic as for the unique case of the US-American Apollo program realised by the *National Aeronautics and Space Administration (NASA)* in the 1960s and early 1970s. The research field of *planetary geology* was born in such a human-robotic context, by applying Earth-bound principles (see Chap. 2), and matching them with remote-sensing measurements, and by returning their data to Earth for analysis. The main tool for observing and analysing extraterrestrial surfaces is *remote sensing* which uses a variety of platforms, in addition (Fig. 3.1) to Earth-based or astronomical

S. van Gasselt (✉)
National Chengchi University, No. 64, Sec 2, ZhiNan Rd., Wenshan District, Taipei 11605, Taiwan
e-mail: svg@nccu.edu.tw

A.P. Rossi
Jacobs University Bremen, Campus Ring 1, 29795 Bremen, Germany
e-mail: an.rossi@jacobs-university.de

D. Loizeau
Université Lyon 1, Lyon, France
e-mail: damien.loizeau@univ-lyon1.fr

M. d'Amore
German Aerospace Center (DLR), Berlin, Germany
e-mail: mario.damore@dlr.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_3

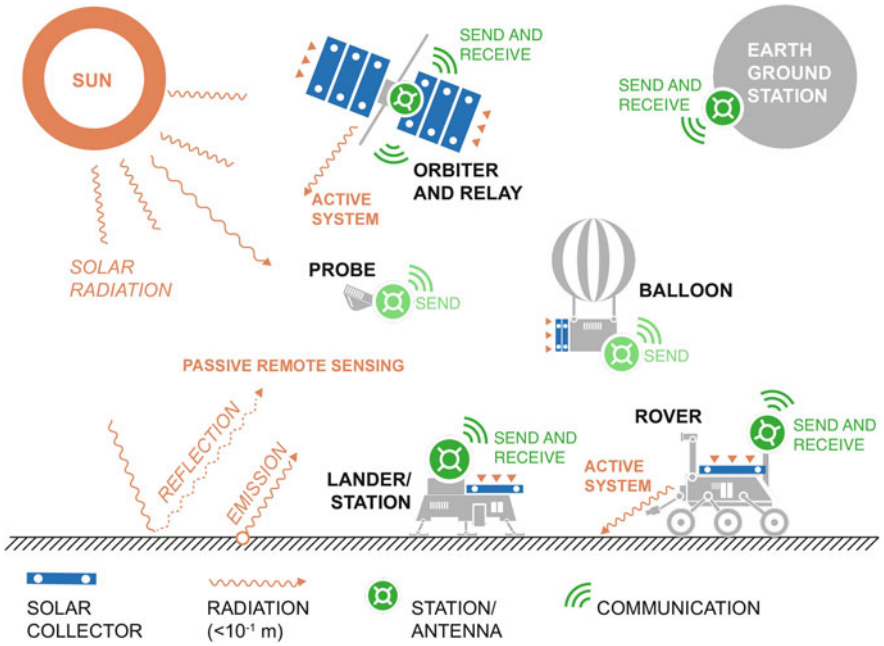

Fig. 3.1 Cartoon depicting main platforms involved in planetary exploration. On most platforms remote sensing experiments can be accommodated. On surface platforms (landers, rovers) also geological in-situ instruments and small laboratories for analysis can be hosted

observations. The basic principles and geological applications of remote sensing are introduced in Sect. 3.2.

Remote sensing is a methodology to gather qualitative or quantitative information of an object from distance and without physical contact, usually through information transported by electromagnetic radiation from instruments on orbiting platforms. Remote sensing data are usually characterized by the geometric and spectral resolution of the observation, the observation time and scientific contents (Sect. 3.2). Historically, remote sensing observations have been the first to be made to explore planetary objects and for most bodies in the Solar System such observations have remained the only ones. In few lucky cases—as for the Moon, Mars, few asteroids and comets—additional data came from samples recovered by robotic probes or humans, or they were delivered to Earth as meteorites (Sects. 6 and 7). With few notable exceptions, the latter usually miss contextual geological information, e.g. exact source location or provenance of a certain meteorite from within a large, geologically complex body.

The ultimate goal in the exploration of extraterrestrial environments is human, but the only attempt thus far was targeted at the Moon in the 1960s and early 1970s in the context of the Apollo program. A renewed geological exploration of the Moon might well start soon, while other targets, such as near-Earth asteroids (Chap. 13) or

Table 3.1 Platforms for planetary exploration and Solar System bodies visited

Type	Moon	Mars	Venus	Mercury	OSS	SSB
Flyby	6	5	10	1	3	2
Entry probe	0	0	5	0	1	0
Orbiter	27	14	7	1	2	3
Impactor	11	0	0	0	0	1
Lander	9	4	7	0	0	2
Rover	3	5	0	0	0	0
Sample return	8	0	0	0	0	1
Human	6	0	0	0	0	0

OSS Outer Solar System objects, *SSB* Small Solar System Bodies.
Source: NASA NSSDC

Mars (Chap. 11), require considerably more time. Table 3.1 summarizes the state of the art of planetary exploration through available platforms.

Robotic exploration has many assets: it can be carried out essentially in each environment by sending a spacecraft (or, in the case of Earth, an aircraft, drone, robot or alike), it can typically record, retain and send back data for analysis using a wide range of experiments once commands are implemented. However, remote sensing approaches when performed exclusively are often affected by considerable ambiguity in absence of ground truth. Local ground validation exists on the Moon with both landers/rovers and samples and on Mars with ander/rovers. Main limitations of in-situ measurements for planetary applications when compared to terrestrial ones are in terms of mass, power, sensibility, overall operating conditions, or data rates.

Remote sensing experiments can be classified in different ways. The most common one is based on where and how the electromagnetic radiation is generated to investigate the object from far. While *passive remote sensing* takes advantage of either radiation generated by the target body or by external sources such as the Sun (e.g., X-ray radiation, infrared radiation, microwaves, see Fig. 3.2) which interacts with the target body, *active remote sensing* is based on electromagnetic radiation generated by the spacecraft experiment, which then interacts with the target body (e.g. radar, laser).

3.2 Imaging

Imaging remote-sensing instruments, such as classical analog or modern digital camera systems employed on orbiting and roving platforms operate in the visible wavelength from about 390–700 nm (0.39–0.7 μm) to the near infrared with wavelength of up to 900 nm. Due to their close resemblance to what the human eyes perceives imaging data are the most accessible means to study planetary surfaces and their features. Classical photographs and digital images enable us to investigate

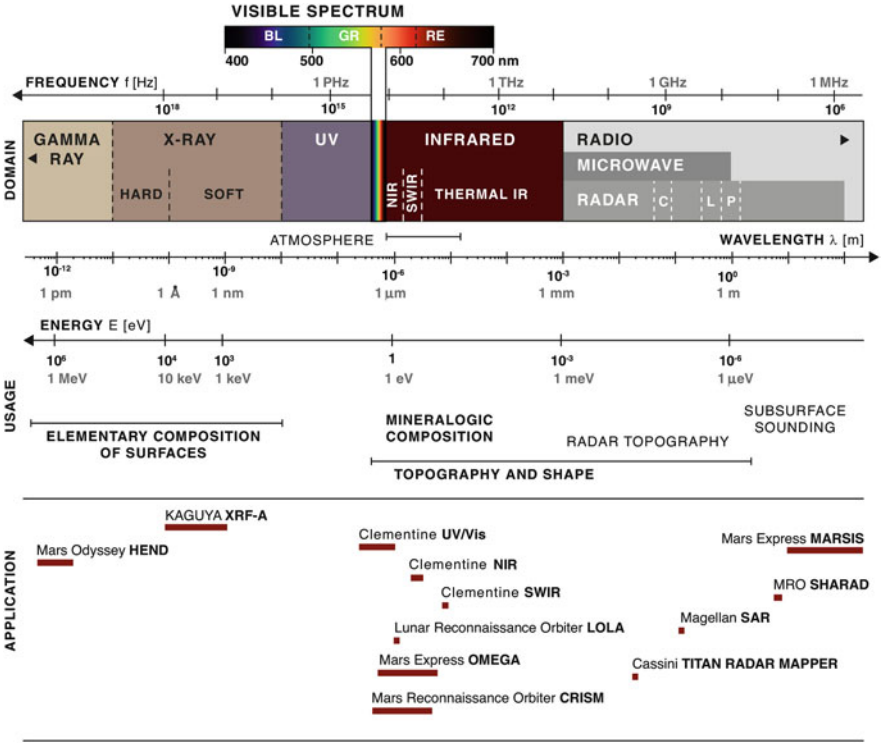

Fig. 3.2 Regions of the electromagnetic spectrum as used in planetary remote sensing and examples of applications on past and recent missions

morphologic characteristics such as relative positions, extents, sizes, structures, textures and colours of surface units and they allow us to eventually transfer these pieces of information into thematical maps. In order to study the actual composition of surface units, visual inspection of image data does not provide any direct clue and a variety of different spectrometers are employed which provide compositional details at wavelengths beyond the sensitivity of the human eye.

All imaging instruments consist of a radiation collecting and focusing unit, such as a telescope, and a unit which converts and stores the brightness pattern retrieved at the sensor on analog film or digital storage medium. Data, either digital or analog, are more than just pictures to look at if the instrument is properly calibrated and scientific data processing has been undertaken. Each grain on a photographic film and each pixel on a digital sensor carries physical information about the interaction between radiation and the surface. Due to calibration, we can directly measure the amount of light that falls onto the sensor at the spacecraft I_1 as so-called *radiance* L $[\mathrm{W} \cdot \mathrm{sr}^{-1} \cdot \mathrm{m}^{-2} \cdot \lambda^{-1}]$.

As we know how much light was sent to the surface it was reflected from and as we know the radiation source usually well, we can relate that quantity of incoming

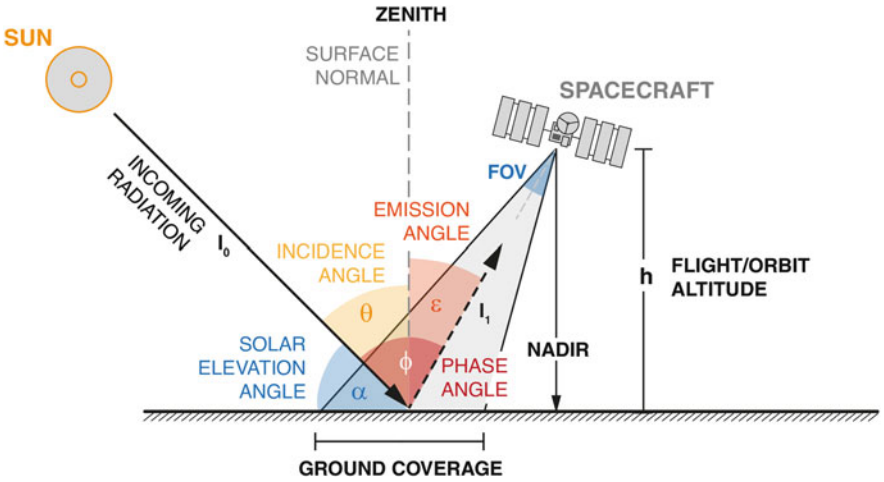

Fig. 3.3 Observation geometries and angles

radiation I_0 to the quantity collected at the sensor I_1 and derive the dimensionless so-called *reflectance* $\varrho = I_1/I_0$ [–].

$$\varrho = \frac{\pi \cdot \mathrm{L}_\lambda \cdot \mathrm{d}^2}{\mathrm{E} \cdot \cos\theta}[-] \tag{3.1}$$

Here, L_λ is the (wavelength-dependent) spectral radiance, d is the distance between radiating object and surface, E is the solar constant and θ is the solar zenith (or incidence) angle, i.e. the angle between the center of the Sun's disk and zenith. (see Fig. 3.3). The solar zenith angle is complementary to the solar elevation angle α. The emission angle ε describes the angle between outgoing (reflected or emitted) radiation and zenith, and the phase angle φ describes the angle between the Sun and the observer, i.e. the instrument.

These basic principles and geometric properties remain essentially the same for all imaging devices and spectrometers, however, the way information is collected and stored has significantly changed over the last 50 years of planetary remote sensing. Imaging cameras covering the visible wavelength range have always been playing a pivotal role for retrieving the first-ever impression of an unknown object, for flyby reconnaissance, or for systematic studying surfaces from orbiting platforms or from landers and rovers. On Earth, photographic film was widely used to image surfaces from spacecraft starting in the 1960s and 1970s during the US Keyhole program. While on Earth exposed film could be sent back by dropping it to the Earth's surface and catching it in-flight, the retrieval of analog film from platforms orbiting planets posed a major challenge. Early Soviet and US-American programs employed photographic cameras on lunar missions with great success by making use of an automatic film processing and scan unit. The

Soviet Yenisey-2 photo-television system operated on Luna 3 (1959) and took the first-ever image of the lunar farside, while the US Lunar Orbiter program (1966–1967) employed a similar system for systematic mapping and landing-site reconnaissance for the Apollo program. Both units operated essentially in a similar way by automatically photographing the lunar surface and by developing exposed film in a complex processing system onboard the spacecraft. After the developing and drying process, photos were automatically scanned and subsequently recorded on a television system which transmitted the modulated electromagnetic signal back to Earth where it was played back from a TV and recorded on film again.

Due to its complexity in implementation and error-proneness the automatic film-development system could not establish and the US as well as the USSR switched to the Vidicon system which was very successfully employed in particular during the Mariner and Viking missions to Mars. A phototube—basically a reversed TV-tube—lets incoming radiation pass through a vacuum tube and temporarily create an electric resistance pattern which is proportional to the amount of incoming light, i.e. brightness, on a sensitive surface. That pattern was scanned line-by-line, transmitted to Earth and wiped out again on the sensor to make room for the next recording. Vidicon tubes were often relatively heavy and prone to errors whenever the previous signal could not be completely erased, causing ghosts artifacts. Also, the tube design caused severe radial distortions of the image so that so-called *reseau marks* (small crosses covering the optics) were needed for later image rectification.

In the 1990s finally, digital camera technology was also applied in planetary exploration by making use of *Charged-Coupled Devices* (CCD) and *Complementary Metal Oxide Semiconductor* (CMOS) detectors which can be found today in basically all digital consumer-market cameras. These detectors allow for an improved handling of data and provided higher resolution, quality and integrity.

Along with the change of detectors and smaller sizes of CCD picture elements (pixels) (usually in the range of few μm), telescopes have been improved and allowed for higher geometric resolution that are determined by focal length and instrument aperture. In the late 1960s, the Mariner 4 camera system produced images at scales of 300–3000 m/px, in the late 1970s the Vidicon-based Viking Imaging System (VIS) allowed for ground-pixel sizes of up to $<$10 m although most image data were collected at spatial resolutions of several tens to hundreds meters per pixel. They were eventually transferred to the global *Mars Digital Image Mosaic* (MDIM) with a resolution of 231 m/px (or 1/128 px/degree). The Mars Orbiter Camera (MOC, 1996–2006) obtained resolutions of up to 3 m/px and in 2010 the Mars Reconnaissance Orbiter HiRISE camera could get down to 30 cm pixel sizes. The same picture is drawn in lunar exploration and translates to an increase in resolution of roughly one magnitude per decade.

The possibilities of improving spatial and also spectral resolution however come with a limitation regarding the amount of data that can be transferred from the sensor to onboard-storage and through interplanetary space back to Earth which is related to the spacecraft's power budget—and therefore performance—of telemetric systems. Also, the need for higher resolution imaging required more sophisticated approaches for the design of sensors. While imaging until the late 1990s was

characterised by *frame-camera* systems, allowing to take (mostly quadratic) fixed-size surface snapshots as a whole at a given instance of time, modern systems have been making use of *pushbroom* and *pushframe* technology. Here, the surface is imaged line-by-line, similar to a scanning device. With these new sensors, long image strips can be acquired while the spacecraft progresses in orbit over the surface. However, such an advantage comes at the cost of simplicity as data processing becomes more complex and observation and illumination geometries as well as time changes from one scanline to the next. It is noteworthy that the only so-called *whiskbroom* system (side-way scan) ever employed on a planetary orbiting spacecraft was the optomechanical scanner of the late Apollo 15–17 mapping cameras that imaged the lunar surface from orbit in a similar way the Earth observing Landsat satellite and its descendants have been operating since 1972.

3.3 Composition and Properties

The study of the composition of planetary surfaces requires different sensors that are sensitive to various types of radiation and wavelength ranges to record radiation from the surface of the body. This radiation can take different forms:

- solar electromagnetic radiation reflected from a surface (in the UV, visible domain or near IR),
- electromagnetic radiation emitted by the surface (in the thermal IR domain),
- electromagnetic radiation emitted by an instrument onboard and reflected by the surface (in the radio domain, or in the near IR),
- high energy radiation (gamma rays or neutrons) produced by the interaction of the surface or near subsurface with cosmic ray particles.

Radiation interacts with the bonds between atoms inside the crystals and absorbs radiation preferentially at given wavelengths. However, when the planetary body is surrounded by an atmosphere, radiation goes through it before reaching the orbiting instrument, and also interacts with the molecules in the atmosphere, and can eventually even be absorbed before reaching the ground.

The infrared (IR) portion of the EM spectrum is of particular interest in Geology as most rock-forming minerals have typical responses. Some instruments in the UV to IR work as single point spectrometers, others are spectral-imagers. Spectral imagers take images line by line or point by point and record portions of the radiation spectra for each pixel of the image. Thus, a single spatial location relates to spectral characteristics taken at different wavelengths.

Infrared wavelengths (Fig. 3.2) are suitable for remote identification of several important rock-forming minerals. Near-infrared (NIR, 0.75–1.4 μm) and short-wavelength infrared (SWIR, 1.4–3 μm), as well as mid infrared (MIR, 3–8 μm) record specific adsorption bands for many silicates. This applies also to hydrated minerals, including silicate hydrated ones (e.g. on Mars, and in Fig. 3.4). Some minerals though, notably quartz or feldspar, lack diagnostic features in these spectral

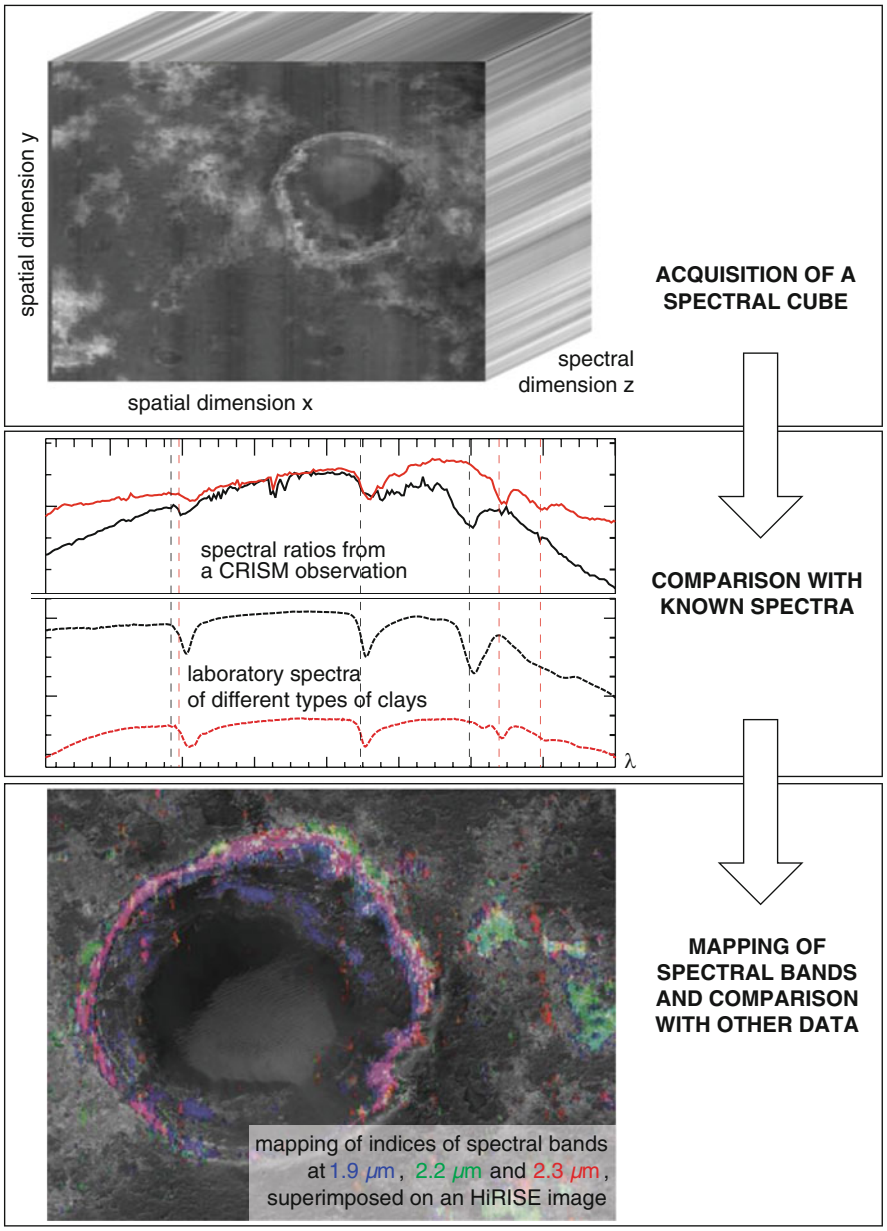

Fig. 3.4 A spectral cube and the main steps of its analysis. Source: Images NASA/JPL/University of Arizona; Laboratory spectra extracted from the RELAB Spectral Database; Copyright 2014, Brown University, Providence, RI

ranges. Thermal infrared (8–15 μm) spectral ranges can help identifying some of those minerals not properly addressed by VNIR/SWIR. Non-imaging spectrometers in TIR range were used successfully on Mars respectively on orbiter and rover platforms: TES, miniTES; other imaging TIR experiments are going to observe (in the 2020s) Mercury (MERTIS).

Geological infrared imaging can also benefit from experiments designed for atmospheric measurements mainly, such as VEX VIRTIS, taking advantage of few *atmospheric windows* where signals from the surface can be detected, such as anomalously hot portions of the surface, possibly also matched by visible brightness increase (Chap. 8), e.g. linked to possible recent or present volcanism.

The final data product created by hyperspectral imagers are spectral cubes (see Fig. 3.4), i.e. stacks of images, each corresponding to a different wavelength.

Ultraviolet (UV) spectral wavelengths (Fig. 3.2) are not very used in geological remote sensing, but they have been used for observing planetary atmospheres for typical gas adsorptions in such spectral range. Indirect geological applications could be, so far only Earth-bound, on volcanic gas emissions, for example. There are several specialized techniques, relying on different portions of the electromagnetic spectrum, that are not mentioned in this chapter (see Fig. 3.2).

Gamma and X-ray wavelengths (Fig. 3.2) in planetary remote sensing applications are used for identifying elemental composition of the uppermost layers of a planetary surface. Due to atmospheric absorptions they can only be used on planetary objects lacking an atmosphere or on the ground when in close or direct contact with a sample. Orbital applications rely on the interaction of solar radiation with the surface while surface application usually require active remote-sensing approaches. The case of high-energy radiation involves detectors that are poorly collimated, the detection is made for the whole area under the spacecraft (or within a solid angle below the spacecraft, intersecting the planetary surface), which depends on its elevation. Hence, mapping of hydrogen and other atoms is performed at different scale when compared to previous methods. Refinements of mapping can be made after many orbits, and methods of collimation are developed for the new instruments. In general, data derived from such experiments are of relatively low spatial resolution, compared to imaging experiments, due to the need of long integration times over multiple passes of the spacecraft over the same region. Therefore large-scale or global elemental maps from Gamma and X-Ray spectroscopy may take several years of observations to be produced by orbiters. Examples of such experiments are the Gamma-Ray Spectrometer (GRS) suites on board Mars Odyssey and LEND on board Lunar Reconnaissance Orbiter.

After recording a spectroscopic signal from the surface, one needs to compare it to known spectra, recorded in a laboratory for example, with well characterized materials. This will enable detecting, mapping and possibly quantifying the presence of atoms or minerals at the surface (Fig. 3.4). In the case of IR spectroscopy this reference is usually contributed by laboratory spectra. In the case of high-energy spectroscopy in the X-ray domain the reference is usually established by carrying sample material of known composition which is exposed to the Sun during mapping.

At the scales we observe (from few tens of meters per pixel to few hundreds of kilometers), surfaces are heterogeneous with different rocks and different types of regolith, and even at microscopic scale, rocks are heterogeneous, so a single spectra, unless taken in a laboratory using a controlled sample, is always a mixture of different minerals and different autochthonous and allochthonous components, such as bedrock and dust.

Here are some recent examples of instruments that mapped the composition (atomic composition or mineralogy) of bodies of the solar system, the list is not comprehensive but tries to cover instruments with major contributions on a diversity of bodies:

- Mars
  - Thermal Emission Spectrometer (TES) on Mars Global Surveyor (MGS),
  - Observatoire pour la Minéralogie, l'Eau, les Glaces et l'Activite [Visible and Infrared Mineralogical Mapping Spectrometer (OMEGA)] on Mars Express,
  - Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) on Mars Reconnaissance Orbiter (MRO),
  - Gamma-Ray Spectrometer (GRS) on Mars Odyssey (MO),
- the Moon
  - Moon Mineralogy Mapper (M^3) on Chandrayaan-1,
  - Lunar Exploration Neutron Detector (LEND) on Lunar Reconnaissance Orbiter (LRO)
- Vesta and Ceres
  - Visible and Infrared Spectrometer (VIR) on DAWN,
  - Gamma Ray and Neutron Detector (GRaND) on DAWN
- P67/Churyumov-Gerasimenko
  - Visible and Infrared Thermal Imaging Spectrometer (VIRTIS) on Rosetta

3.4 Topography and Structure

Deriving information about topography, together with imaging, is one of the key tools used to remotely study the geology of Earth or any solid extraterrestrial body. There are several ways to construct or extract topographic models of a planetary surface:

- stereo-photogrammetry using passive optical sensors and data (e.g. taking advantage of Solar radiation or emission from target bodies),
- laser scanning using active optical sensors by measuring return times of light signals sent to the planet's surface,

- radar imaging using active sensors in the microwave/radar domain allowing to make observation through thick atmospheres such as present on Venus or Saturn's moon Titan.

The result of any technique might be a local, regional or global topographic model (Chap. 11).

Information about topography itself is of importance not only to describe and understand the geomorphology but also to reconstruct the geology of any planet or satellite. The combination with imaging, compositional and contextual information are synergistic for geological interpretation as topography allows to reconstruct the relative and absolute position of geologic units and their extent. In each snapshot the topographic signature of present or past impacts (Chap. 7), endogenic (Chap. 8) and surface (Chap. 9) processes and overall geodynamics (Chap. 10) are recorded. The hypsography, i.e., the cumulative histogram of the topography of a planetary body, summarizes at a glance the overall distribution of heights, highlighting, for example, the large amount of volcanic plains on Venus, the oceanic vs. continental lithosphere on Earth, Mars' dichotomy or the primary and secondary crusts of the Moon, respectively anorthosite highlands and basaltic maria (Chap. 11).

Photogrammetric restitution of the surface topography is based on the calculation of viewing geometry, typically imaging the same portion of a planetary surface under different angles (Fig. 3.3). It is based on the basic principles introduced in Sect. 3.2 and it takes advantage of the parallax between two or more cameras pointing at the same area/object in order to derive the distance and, with some additional or a-priori knowledge, the absolute height/topography. Essential to its application are detailed knowledge about the interior information of a camera system and the observation geometry. Photogrammetry is the oldest technique for topographic creation, after trigonometric land-based surveying.

Alternatively, coherent laser light pulses are used for deriving information about surface topography. Once a laser pule is sent off from an altimeter system, its time to return is used to calculate the distance to the surface and back to the detector. In the case of Mars Global Suveyor's Mars Orbiter Laser Altimeter (MOLA) during 1996 and 2006 the surface footprint of a single laser shot was about 160 m in size due a *laser beam divergence* of 420 μrad and an orbit altitude of roughly 378 km. For the Lunar Reconnaissance Orbiter (LRO) Lunar Orbiter Laser Altimeter (LOLA), the beam divergence was reduced to 100 μrad and orbit altitude was significantly lowered to 50 km resulting in footprints of 5 m diameter. Also, the frequency with which laser pulses have been sent to the planet's surface were improved from 10 Hz for the MOLA instrument to 28 Hz for LOLA resulting in a finer along-track spacing of measurements. Orbital laser altimeters and scanners sample the surface along tracks with a dense point spacing *along track* and considerable point spacing *across track* depending on the exact orbit geometry. Naturally, in a polar orbit point densities are higher near polar areas when compared to equatorial areas. To overcome this, point data need to be interpolated to be able to provide gridded maps with continuous data coverage.

The behaviour of most geological materials at radar wavelengths (Fig. 3.2) provide ways of distinguishing them, although usually with less detail compared

to optical remote sensing depending largely on the combined effect of surface dielectric constant and surface roughness scaled to radar wavelength. Radar imaging has been used extensively in studying the surface of Solar System bodies with thick atmospheres, such as Venus (Magellan) and Titan (Cassini-Huygens). Atmospheres can be partially or totally opaque at visible wavelengths as well as in other portion of the visible spectrum. However, the potential of radar imaging and mapping has not been exploited yet for planetary exploration and they are mainly used for either imaging purposes or for the reconstruction of topography using single radar pulses in a somewhat simpler process compared to radar image reconstruction.

Single footprints of radar echoes sent to the surface of a planet or satellite are typically relatively large as they are defined by the size of the antenna. In order to improve spatial resolution the antenna's size is *synthetically* increased by making use of the spacecraft's motion in orbit and by processing the signal returns with respect to their specific time. This method, termed *Synthetic Aperture Radar (SAR)*, allows for much higher spatial resolutions without physically creating larger antennas. In the case of NASA's Magellan mission to Venus, radar altimeter footprints typically ranged between 10 and 30 km (e.g. for a spacecraft orbiting down to $\sim$300 km of altitude). Very high resolution topography can be obtained from SAR data using radar interferometry, e.g. SRTM is a notable example from the 90s when the Space Shuttle, equipped with radar antennas, collected SAR data that were later processed to produce quasi-global topography. SAR interferometric measurements have not been performed on planets other than Earth so far.

3.5 Geophysical Tools

Geophysical probing of Earth's near and deep subsurface relies on a variety of methods, using several parts of the electromagnetic spectrum, either remotely, in-situ or in close contact. A very wide range of geophysical exploration techniques have been developed through time on Earth, driven by resources (oil and gas, minerals) in most cases. Among those available and used for terrestrial applications, a fraction is used for planetary geological exploration. Similarly to what applies to imaging remote sensing, some geophysical techniques rely on passive observations, other on active ones, the latter being much less used, due to their inherent higher operational complexity on extraterrestrial settings.

3.5.1 Potential Fields

Gravity and magnetic field measurements are most used for investigating the physical state and subsurface structure of planets and moons, as well as typically the first and most affordable large-scale surveying options for terrestrial oil and mineral exploration (e.g. from satellite or aircraft platforms). Gravity variations, or

anomalies (which tend to be the most used in geology, more than absolute values) are sensitive to the mass distribution inside a planet, this can have close correlation with surface and subsurface geology, particularly in its topmost layers.

The detail and spatial resolution at which gravity data can be directly measured is strongly dependent on the distance between measurement and the size/scale of the source, thus the orbit altitude of the platform. Apart from direct measurements in-situ with gravimeters, one could use combined topography, spacecraft orbit data, and spacecraft tracking information (radio science). Global gravity data are often used in conjunction with topography in order to construct models of the subsurface, such as crustal or lithospheric thickness (see Chap. 8).

Magnetic field measurements, especially if carried out far from planet or moon surfaces can carry variable information on the subsurface. Magnetic field values can be measured using specific experiments (magnetometers) or can also, in some limited cases, derived from other types of measurements (e.g. ionosphere-sounding radars, such as MARSIS). Magnetic measurements allow for detecting not only currently present global magnetic fields (e.g. on Earth), but also ancient ones, whose record is preserved as remnant magnetization in minerals within the crust, still measurable from orbit, like in the case of Mars. In any case, the presence of a global interior magnetic field is strongly linked to the interior state of any planet or moon (Chap. 10).

3.5.2 Seismics and Subsurface Sounding

Both active and passive acoustic probing of the subsurface are well-established techniques on Earth used since many decades and with a large variety of applications. They make use of generation, transmission, reflection and refraction of seismic waves through Earth's crust in order to derive information on the physical characteristics and—indirectly—of its geological nature, such as composition, lithology, state of fracturation. In such cases this is achieved through *ground truth*, which for the subsurface typically consists of drilling or coring. Also, the combined use of independent geophysical techniques (e.g. acoustics and electromagnetics) allows for better constraining the in-principle unknown or poorly known subsurface geology. In the planetary case, the subsurface is largely unknown in absence of a subsurface sounding experiment and very often contextual or *ground truth* information is missing. Seismics on solid bodies other than Earth have been successfully performed only on the Moon (see Fig. 3.5). Active seismic investigations on the Moon were performed in later Apollo missions.

Seismic waves can be produced actively through percussion (such as hammering) or explosions (such as by mortars used during Apollo), while passive seismics could make use of earthquakes (moonquakes) or possible energization due to small asteroid impacts, particularly on atmosphere-less bodies such as the Moon, but also on planets with a relatively thin atmosphere, such as Mars.

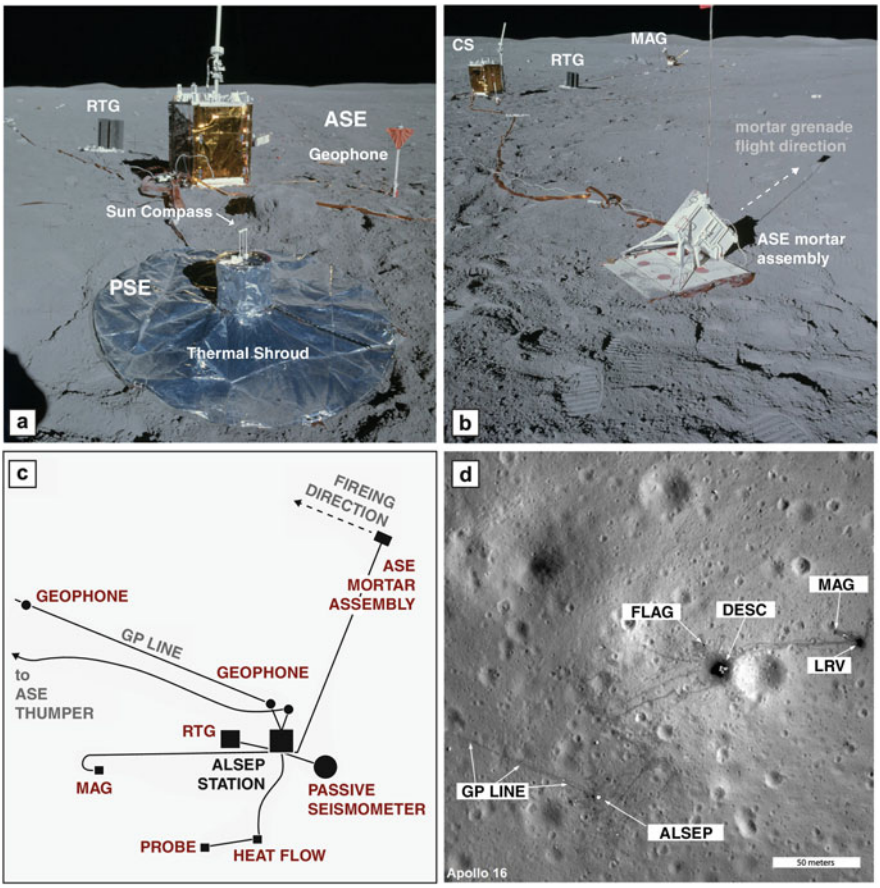

Fig. 3.5 Seismic experiments on the Moon. (**a**) The Apollo 16 Passive Seismic Experiment, PSE; (**b**) the Apollo 16 mortar package of the Active Seismic Experiment, ASE; (**c**) overview of the ALSEP surroundings with ASE and PSE equipment, cf. (**a**) for reference; (**d**) high-resolution image of the AS 16 landing site with ALSEP location and larger context. Label ALSEP refers to the Apollo Lunar Surface Experiment Package, ASE and PSE refer to the Active and Passive Seismic Experiment, respectively. RTG refers to the radioisotope generator, MAG is the surface magnetometer, CS is the ALSEP Central Station. GP labels a Geophone and DESC refers to the Apollo Descent stage. Source: (**a**), (**b**) AS16-113-18347. (**d**) from NASA/LRO landing site montage, 23 Nov 2015

Alternatively once more radar wavelengths can be utilized to investigate the subsurface of planetary bodies. Subsurface sounding radars were so far employed only on the Moon (ALSE) during Apollo times, and later by JAXA's Selene/Kaguya (LRS) on Mars twice, with MARSIS and SHARAD on board ESA's Mars Express and NASA's Mars Reconnaissance Orbiter (MRO), respectively. The fact that water ice is transparent to a wide range of radar wavelengths (Fig. 3.2) makes radar sounding ideal to investigate the subsurface of icy or ice-rich bodies (glaciers,

planetary polar caps, permafrost or ice-rich regolith). On the other hand liquid water, due to its very high dielectric constant, is preventing radar signals from crossing water-saturated levels.

Future subsurface exploration of icy satellites in Jupiter and Saturn systems will also rely on orbital radar sounders able to penetrate the icy crust (see Chap. 12). Ground-based radar sounding is very common on Earth (GPR) and are also bound to planetary targets (Mars) on board upcoming rover missions (ExoMars). A more detailed treatment on planetary interiors, investigated through measurements and models, is provided in Chap. 10.

3.6 Landing Sites and In-Situ Tools

Going beyond remote sensing on any planetary surface requires landing somewhere in order to perform experiments. The selection of an adequate landing sites is extremely important for both in-situ exploration and, even more, for sample return. Landing sites are literally the locations for *ground truth* for extraterrestrial bodies (Chap. 5).

Any landing site has to satisfy scientific and engineering requirements, including safety, which can vary substantially depending on the mission scenario and the knowledge of the target planetary body: the selection of landing sites is a complex process and constrained by available data and information. In the case of very remote and complex landings, such as Huygens on Titan, the choice of an exact landing site is limited, therefore spacecraft need to be designed for broad range of survivability options. Early Soviet probes to Venus but also the modern Huygens lander had floating capabilities as it was not known whether the landing sites were liquid or solid surface. In most recent cases for Mars landing sites have been selected based on an increasingly large base of data used to take both engineering and scientific decisions. In some cases, exact landing sites are being selected with the help of instruments on associated orbital platforms and direct observations. In that way, the landing sites for the Viking 1 and 2 landers were selected and characterized as much as possible from orbit by the Viking Orbiter instrument suite before the landers were actually deployed. Details on how planetary *ground truth* is achieved are provided in Chap. 5.

Planetary surfaces visited by landers or rovers can be imaged through close-range remote sensing, but a wide set of in-situ analytical capabilities exist, ranging from small contact experiments to relatively large (but typically very small for Earth standards) in-situ analytical laboratories.

3.6.1 Contact Experiments

Remote sensing-based analyses can provide data on mineralogical or chemical composition, but more detailed local studies require in-situ approaches through

different techniques and experiments. Those can be mounted on robotic arms for easy access to the spot of investigation (Chap. 5) or hosted elsewhere on landers and rovers and fed by robotic appendices (arms, scoops, etc.).

Rover-based in-situ analytics for the determination of compositions can include some of the following instruments.

- Alpha Particle X-ray Spectrometry (APXS) generates x-ray excitation irradiating sounded materials with α particles produced by radioactive decay. Collected spectra are diagnostic of the chemical composition of the analysed rock.
- X-Ray diffractometry, widely used in geology and mineralogy for terrestrial applications was also used on Mars (CheMin on MSL): through irradiation of samples with X-Ray that are diffracted by the mineral structure, producing indicative spectra.
- Mössbauer spectrometry uses γ-Ray in order to provide mineralogical compositional information as well. It requires materials to bear Iron.
- Laser is used to investigate materials with Raman spectrometry: it can be used to identify both minerals and organic materials, thus, it is of interest for exobiology research (Chap. 14).
- LIBS experiments measure spectra indicative of elemental composition by collecting spectra of surface materials vaporized by laser pulses. A LIBS experiment, ChemCam, is currently operated by Curiosity (MSL) on Mars

Several variations of the above mentioned analytical techniques exist and they have either already been used on past lander or rover missions or they are going to be developed for upcoming ones.

3.6.2 In-Situ Laboratories

Automated acquisition, processing and analysis of extraterrestrial samples taken from the surface or sub-surface can be carried out with instruments and handling subsystems onboard landers or rovers. Such automated sample analyses have been achieved thus far on Mars, their advantage is in the much more time- and cost-effective handling of samples without the need of returning them to the Earth. Compared to other in-situ instruments presented in the previous section, the advantage is to better control the conditions of observation: light, temperature, air, solvents, liquid or gas reagents, size and state of the sample. The instruments can also be larger than those mounted at the tip of a robotic arm. This allows for more precise and complete studies of samples than with contact instruments as one can take microscopic images, test the physical or biological behaviour of the sample, identify minor compounds, differentiate isotopes, or identify organic molecules, and quantify these species. On the other hand, one needs to take the samples from the studied body, process them, and bring them to the instrument, which adds significant complexity to the mission.

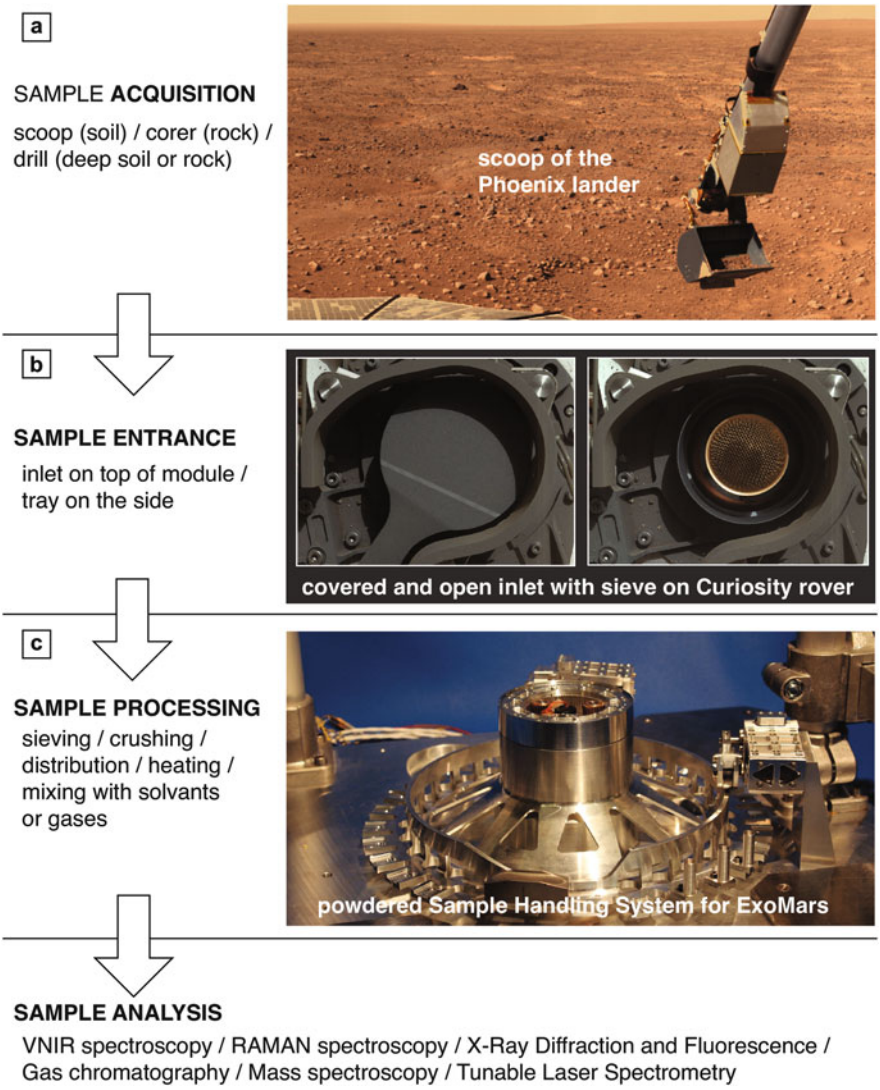

Fig. 3.6 Series of steps needed for sample analyses by in-situ laboratories. Source: (**a**) NASA/JPL-Caltech/university of Arizona. (**b**) NASA/JPL-Caltech/MSSS. (**c**) OHBSystem AG, 2016, courtesy of the ExoMars Project

For sampling, a variety of devices have been used, depending on the target and the type of lander or rover (Fig. 3.6). For atmospheric sampling, pumps such as for the Huygens lander during its descent through Titan's atmosphere have been used. Alternatively, atmospheric gas can be sampled by filling a collector passively through valves. First surface sampling techniques involved a scoop at the end of

an arm, taking samples of soil or regolith and bringing them to the laboratories on-board the probe. The technique works only with loose material, and was used on Mars on the Viking and Phoenix landers. Scoops were also employed on the autonomous Surveyor landers to the Moon in the 1960s. The main purpose here was to image the material and assess its physical properties using photographs. For harder material like rocks, drilling is needed. It also allows to obtain samples from greater depths. Drills of different lengths have been used or are planned on the Venera 13 and 14 landers on Venus, on the Curiosity and ExoMars rovers on Mars, and on the Philae lander on comet 67P.

Sample processing can involve sieving in the case of soils and regolith, or crushing it in the case of rock drills, and then dosing the samples for delivery to a display unit for the instruments. The display unit varies with the number of samples to be analysed and the instrument. In the case of Martian rovers, with several samples to be analysed, the sample handling an distribution device can be a carousel with cups or cells which delivers the sample to the various instruments, or ovens when heating is needed. Some systems have a limited number of uses (the number of heated samples will correspond to the number of ovens), while others can be used multiple times on different samples.

One can distinguish two broad groups of instruments: those which analyse samples without destroying them, i.e. *non-destructive* analysis (e.g. IR spectroscopy, Raman spectroscopy, X-ray diffraction and fluorescence), and those who need to modify the sample, by mixing the sample with reagents, or separating the different molecules by heating the sample or shooting laser pulses at it. This is as partially or totally *destructive* analysis. Gases expelled when heating can be analysed using tunable laser spectrometry or gas chromatography and/or mass spectrometry, and ions created by laser desorption are analysed through mass spectrometry. In some cases, before heating, the sample can be mixed with chemical solvents to help extracting eventual molecules of interest. Reference standards are also present for calibrating the instruments once on the surface of the planetary body. This is also the case for close-range remote sensing instruments, where colour chips or three-dimensional shapes are present as calibration targets.

In-situ laboratories that studied or attempted to study samples of soils and rocks include:

- Viking 1 and 2, two NASA landers on Mars (Biology and GC/MS experiments)
- Venera 11 to 14, 4 USSR landers on Venus (GC, MS and XRF)
- Phoenix, a NASA lander on Mars (TEGA, MECA)
- Curiosity, a NASA rover on Mars (CheMin, SAM)
- Philae, an ESA lander on comet 67P (Ptolemy, COSAC, Civa)

The ESA ExoMars rover—planned for launch to Mars in 2020—will also carry a suite of internal instruments (MicrOmega, Raman, MOMA). A MicrOmega instrument is also present on the lander MASCOT of the JAXA Hayabusa 2 mission, on its way to the asteroid 1999 JU3, but there it will analyse the samples directly on the ground, without any crushing or processing. NASA's Mars 2020 rover will study

soil and rocks with its mast and arm instruments and then take samples (*caching*) for potential return to Earth years later but will not analyse samples in situ.

3.7 Sample Return

The ultimate step in the mainly robotic scientific exploration of the Solar System is constituted by sample return. Returning samples form extraterrestrial bodies allows to use advanced analytical techniques and sample processing and preparation far beyond what is feasible on any spacecraft. Returning extraterrestrial samples is effort-intensive and costly, though. Therefore the choice of which samples to return is extremely important, but it is not always possible to achieve.

There are different ways to collect samples, depending on the planetary target, its distance to Earth, a priori knowledge and mission duration and profile, as well mass constraints for sample collection and sample return capsule. End members are:

- *Grab-and-go*: i.e. getting all what is possible to get given constraints, as fast as possible, or by
- collecting a well-constrained, carefully selected sample or sample suite.

Examples of grab-and-go for missions to small bodies are JAXA Hayabusa or NASA Startdust to an asteroid and a comet, respectively (Chaps. 6 and 13). Typically, the first steps of selecting and choosing samples is addressed by other exploration platforms and techniques, i.e. either by remote sensing for the selection of landing sites or other instruments providing in-situ measurements.

Even grab-and-go sample return for very short-duration operation (e.g. from an asteroid with quick touch-down like in the case of Hayabusa) requires careful selection of landing/sampling sites, for both safety/engineering and scientific reasons. Well-constrained sampling requires much more complex, multi-scale information both for scientific and engineering purposes, thus much more laborious preparations.

Existing sample return missions have been focused on the Moon mainly (Table 3.1). The other very large source of naturally returned samples is constituted by meteorites (Chap. 6). Moon samples are reasonably well-constrained, although parts of them are ejected material from very large distances (Chap. 7). Asteroid sample return missions have been successfully carried out (e.g., Hayabusa) or are in preparation for launch at the time of writing (e.g., OSIRIS-REx, Hayabusa-2).

Sampling variety (in terms of rock types, scientific objectives addressed) vs. affordable mission execution time and returnable mass always require trade-offs which have to be matched as well with safety constraints. Landing on a very safe spot might be far away from desired target materials, or the safe landing spot might be in geologically homogeneous terrain but the mission requirements include heterogeneous material. There would be a need for roving larger distances, collecting samples and returning them all in a repository (cache) for later retrieval. Both *grab-and-go* and selected, diverse samples share the same problem on planets

with relatively high gravity and an atmosphere such as Mars, where a single mission cannot perform collection, caching and return of samples.

Automated sampling on the Moon (e.g. the Soviet Luna program) followed largely a *grab-and-go* approach. Human robotic-assisted sampling on Apollo missions had more contextual information and mobility, plus geologically trained astronaut judgement that guaranteed sample diversity and—within limits—statistical representativity.

Although most planetary bodies are hostile to life as we know (Chap. 14), both missions and eventual samples recovered from extraterrestrial sources need to be evaluated for *planetary protection*, i.e. the need to prevent forward or backward biological contamination between planetary bodies, missions, samples and the Earth (see Chap. 14). Even astronauts onboard the first few Apollo missions on the Moon had to undergo quarantine, which was eventually lifted for later missions.

Take-Home Messages

Solar System geological exploration uses a variety of techniques and except for the Moon investigations were mostly based on remote sensing covering the full range of the electromagnetic spectrum.

Imaging and topographic reconstructions provide the first basic set of data for qualitative and quantitative geological and stratigraphic analyses

Multi- and hyper-spectral remote sensing combined with imaging and topography/morphology adds valuable and often unique compositional information

Ground truth on planets and satellites is provided typically by robotic landers and rovers, for the Moon it has also been performed by trained humans

Sample return, associated with geological contextual information and careful sample selection, when and where possible, is the ultimate robotic exploration tool for extraterrestrial bodies. Sampling and sample analysis on Earth is unrenounceable

Further Readings

Cloutis, E.A.: Review article hyperspectral geological remote sensing: evaluation of analytical techniques. Int. J. Remote Sens. **17**(12), 2215–2242 (1996). doi:10.1080/01431169608948770

Crawford, I.A., Anand, M., Cockell, C.S., Falcke, H., Green, D.A., Jaumann, R., Wieczorek, M.A.: Back to the Moon: the scientific rationale for resuming lunar surface exploration. Planet. Space Sci. **74**(1), 3–14 (2012). doi:10.1016/j.pss.2012.06.002

Elachi, C., Van Zyl, J.J.: Introduction to the Physics and Techniques of Remote Sensing. Wiley, New York (2006)

Falkner, P., Peacock, A., Schulz, R.: Instrumentation for planetary exploration missions. Treatise Geophys. **10**, 595–641 (2007). doi:10.1016/B978-044452748-6.00171-1

Floyd, F., Sabins, J.R.: Remote Sensing Principles and Interpretation, p. 449. Freeman, New York (1987)

Ford, J.P., Plaut, J.J., Weitz, C.M., Farr, T.G., Senske, D.A., Stofan, E.R., Michaels, G., Parker, T.J., Fulton, D. (eds.): Guide to Magellan image interpretation. http://history.nasa.gov/JPL-93-24/contents.htm (1993)

Fowler, C.M.R.: The Solid Earth: An Introduction to Global Geophysics. Cambridge University Press, Cambridge (1990)

Gellert, R., et al.: Alpha Particle X-Ray Spectrometer (APXS): Results from Gusev crater and calibration report. J. Geophys. Res. **111**(E2), E02S05 (2006). doi:10.1029/2005JE002555

Grant, J.A., Golombek, M.P., Grotzinger, J.P., et al.: The science process for selecting the landing site for the 2011 Mars Science Laboratory. Planet. Space Sci. **59**, 114–1127 (2011). doi:10.1016/j.pss.2010.06.016

Greeley, R.: Introduction to Planetary Geomorphology. Cambridge University Press, Cambridge (2013)

Gupta, R.P.: Remote Sensing Geology. Springer, Heidelberg (2013)

Lognonné, P., Johnson, C.L.: Planetary seismology. Treatise Geophys. **10**, 65–120 (2015). doi:10.1016/B978-0-444-53802-4.00167-6

Mahaffy, P.R., et al.: The sample analysis at Mars investigation and instrument suite. Space Sci. Rev. **170**(1–4), 401–478 (2012)

McLennan, S.M., Sephton, M.A., Allen, C., Allwood, A.C., Barbieri, R., Beaty, D.W., Boston, P., Carr, M., Grady, M., Grant, J., Heber, V.S., Herd, C.D.K., Hofmann, B., King, P., Mangold, N., Ori, G.G., Rossi, A.P., Raulin, F., Ruff, S.W., Sherwood Lollar, B., Symes, S., Wilson, M.G.: Planning for Mars returned sample science: final report of the MSR end-to-end international science analysis group (E2E-iSAG). Astrobiology **12**(3), 175–230 (2012)

Nakamura, Y., Latham, G.V., Dorman, H.J.: Apollo lunar seismic experiment–final summary. J. Geophys. Res. Solid Earth **87**, A117–A123 (1982)

NASA: Chronology of lunar and planetary exploration. http://nssdc.gsfc.nasa.gov/planetary/chronology.html (2015). Accessed June 2015

Short, N.M., Robinson, J.: The remote sensing tutorial. Goddard Space Flight Center, NASA. http://fas.org/irp/imint/docs/rst/ (1998)

Van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., Hecker, C.A., Bakker, W.H., Noomen, M.F., van der Meijde, M., Carranza, E.J.M., de Smeth, J.B., Woldai, T.: Multi- and hyperspectral geologic remote sensing: a review. Int. J. Appl. Earth Obs. Geoinf. **14**(1), 112–128 (2012). doi:10.1016/j.jag.2011.08.002

Wieczorek, M.A.: The gravity and topography of the terrestrial planets. Treatise Geophys. **10**, 165–206 (2010)

Chapter 4
Cartography Tools

Trent M. Hare, James A. Skinner, Jr., and Randolph L. Kirk

4.1 Introduction

The term cartography is not only defined as the study of and the creation of maps, but also includes the principles of cartography combined with the integration of science, techniques, and aesthetics essential for useful map making. The cartographic principles used for terrestrial map making can also be used for extraterrestrial bodies.

The difference is in the shape or size of the planetary body including orbiting moons or an interplanetary body like a comet or asteroid. Nearly all larger bodies in our solar system have defined geodetic parameters, documented by the *International Astronomical Union (IAU)*, allowing capable cartographic applications, like Geographic Information Systems (GIS) and Remote Sensing (RS) applications, to study these bodies. Tools to support modern planetary cartography have dramatically evolved in the last 50 years. Here we present a summary of modern planetary cartography and the tools used to support it.

Planetary cartography is an essential tool for the planetary research community. No serious extraterrestrial geologic study can be accomplished in the absence of good cartography. To have context (Chap. 2), scientists must be able to precisely identify the location of the surface on the planetary body that they are studying. Likewise, others areas of planetary research are also greatly enhanced by the availability and use of accurate and professional cartographic data products. Cartographic standards such as uniformity of coordinate systems, accurate horizontal and vertical positioning, and mapping methods, scales and schema are vital components to allow data usage across many facilities, including surface exploration by robots and humans. To support this, in 1976, the IAU established a *Working Group* on

T.M. Hare (✉) • J.A. Skinner Jr. • R.L. Kirk
Astrogeology Science Center, U.S. Geological Survey, Flagstaff, AZ, USA
e-mail: thare@usgs.gov; jskinner@usgs.gov; rkirk@usgs.gov

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_4

the *Cartographic Coordinates and Rotational Elements of Planets and Satellites* and reports triennially on the preferred rotation rate, spin axis, prime meridian, and reference surface for planets and satellites.

Besides simple sketches created using the naked eye, mapping of the Moon began, in earnest, in the early 1600s with the invention of the telescope (see Chap. 1). Earth-based telescopes were used until the advent of launched spacecrafts in the mid-1900s. In 1964, the United States Ranger VII spacecraft transmitted 4316 high-resolution telescopic vidicon television camera photos during the last minutes of flight before impacting the Moon. These images represent the first large catalog of planetary remotely-sensed images as taken by a robotic satellite. The term for this, *Remote Sensing*, is commonly defined as collecting imagery at a distance. Using these images and images taken by NASA's five Lunar Orbiter flights (1966–1967) and the Apollo missions (1961–1972), the *United States' Defense Mapping Agency (DMA)*, now part of the *U.S. National Geospatial Intelligence Agency*, were tasked to produce the first large-scale controlled lunar photomosaic and topographic maps (Fig. 4.1) of the Moon.

Without the use of computer-based mapping techniques, cartographers would have manually manipulated hard-copy images to mosaics, termed hand-mosaics. The hand-mosaics were created in local map projections using current control networks such that distortions were minimized across the different regions. Mosaics were then artistically improved using airbrush painting techniques to make a consistent base map without seams and illumination differences.

Modern cartographic tools continue to rapidly advance as computer hardware and novel input devices are created. The production of a planetary map requires making a large number of choices beyond just the creation of the primary base map,

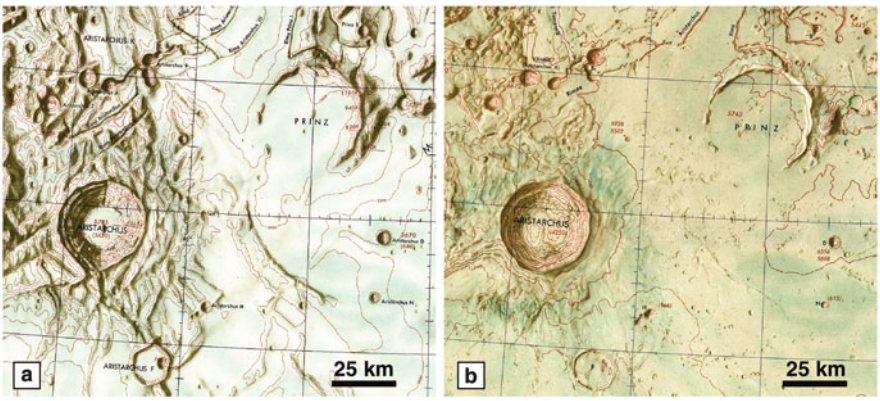

Fig. 4.1 Lunar topographic maps of the Aristarchus region; (**a**) extracted portion of the *Lunar Chart Series (LAC #39)*, scale 1:1,000,000, published by the Aeronautical Chart Information Center, United States Air Force in 1963; (**b**) same region extracted from the *Lunar Map Series (LM #39)*, scale 1:1,000,000, published by the Defense Mapping Agency, Aerospace Center, in 1979. Source: (**a**) NASA, United States Air Force. (**b**) NASA, LPI

such as the details of the coordinate system adopted, the source of geodetic control used to position observations in this system including the map projection, scale, the extent of an individual digital product, attributes and symbologies, metadata, and the placement of planetary nomenclature. Standard tools which address these needs include Remote Sensing (RS) image processing systems and Geographic Information System (GIS) mapping applications.

4.2 Digital Image Processing

Despite the introduction of digital transmitted data in the 1960s, it was not until the late 1980s that digital cartographic processing was technically practical. Today the time-consuming and multi-step manual processes of the past have been superseded by digital cartographic methods. And although much of the process is automated, image mosaic generation is still a complicated process (Fig. 4.2). This stems from fact that the volume of available data has grown dramatically and of the growing need for better positional accuracies. Images are not only just tied to each other but a full adjustment of all the images is essential. The process of adjusting the images as a group is called a bundle adjustment which helps to fit individual images geospatially to known control networks and in turn can create new control networks. Geologic maps are still created by manually drawing of geologic contacts although most use a computer draw-screen, tablet or another input device like a computer mouse.

Most planetary data taken by NASA or other non-US spacecraft are stored in a *Planetary Data System (PDS)* standard. The bulk of the PDS data holdings are cataloged in their original raw instrument form. As described above, to use these

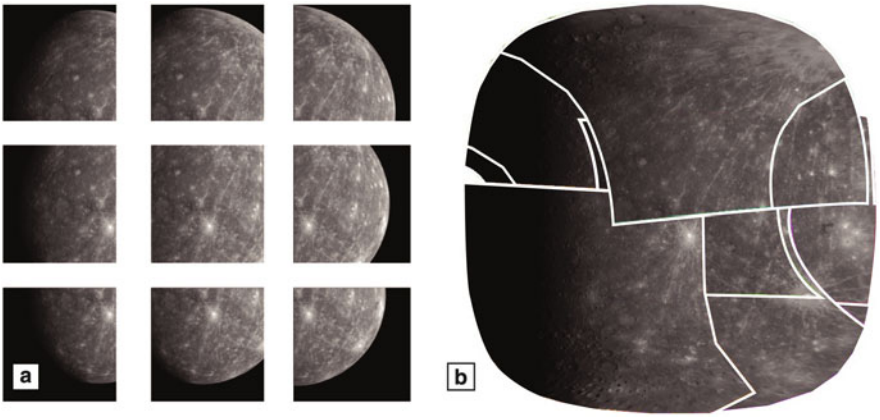

Fig. 4.2 Individual images taken by *NASA's Mercury Dual Imaging System (MDIS)* on the MESSENGER spacecraft (**a**) and an example mosaic (**b**) once the image have been controlled and then merged. Source: T. Becker, USGS

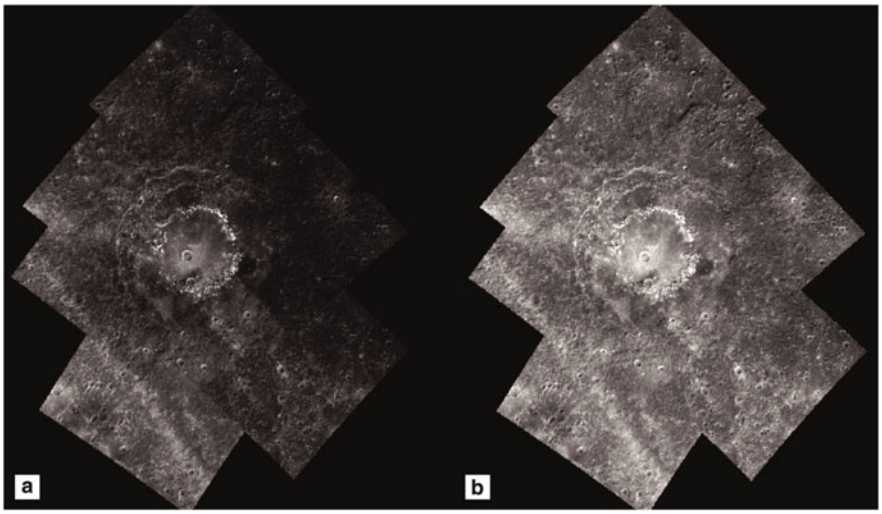

Fig. 4.3 Example image mosaic of Raditladi crater (258 km diameter) from NASA's Mercury Dual Imaging System (MDIS) on the MESSENGER spacecraft showing no photometric correction (**a**) and then the same image mosaic after applying the photometric corrections to the individual images prior to the mosaic creation (**b**). Source: K. Becker, USGS

data sets in cartographic applications they should first be spatially referenced to the planetary body.

Two example image processing systems *ISIS (Integrated Software for Imagers and Spectrometers)*[1] and *VICAR (Video Image Communication and Retrieval)*[2] were created to do just that. Along with geometrical correction, these systems also excel at correcting camera distortions, updating radiometry parameters, performing photometric corrections and combining them into mosaics via bundle adjustment techniques. Radiometric calibration recalculates the values in an image based on exposure time, flat-field observations, dark current observations and other factors describing the unique electronics design and characteristics of an imaging system. Photometric corrections help to adjust images acquired under different illumination and viewing geometries such that the resulting images appear as if obtained under uniform conditions by adjusting the brightness and contrast (Fig. 4.3).

In the early 1990s, ISIS grew directly from the *Planetary Image Cartography System (PICS)*. Both ISIS and PICS were developed by the *Astrogeology Science Center at the United States Geological Survey (USGS)*. To list a few examples, ISIS is currently used to process images from NASA's Mars Odyssey, Mars Global Survey, Mars Reconnaissance Orbiter, Lunar Reconnaissance Orbiter and the Mercury MESSENGER mission.

[1] https://isis.astrogeology.usgs.gov/.

[2] http://www-mipl.jpl.nasa.gov/vicar_open.html.

The VICAR (Video Image Communication and Retrieval) image processing system began in 1966 at the *Jet Propulsion Laboratory (JPL)*. This system processed data from NASA missions like Ranger, Surveyor, Mariner, Viking and Voyager. It has now been upgraded and ported to nearly all modern operating systems and is still used by many facilities to process data for Earth, Mars and other bodies. VICAR is currently being used for NASA's Cassini mission and the European Space Agency's (ESA) High Resolution Stereo Camera (HRSC) aboard the Mars Express spacecraft.

It should be noted that the use of ISIS, VICAR, or other software to transfer image data to a map-projected space does not guarantee that the resulting map coordinates of features are strictly accurate. Positional accuracy depends on the ability to measure or reconstruct the position of the spacecraft and the pointing of the instrument as each observation was taken, as well as knowledge of the shape and rotation of the target body (Fig. 4.4). To assist with this, the *Navigation and Ancillary Information Facility (NAIF)*, housed at JPL, provides an information system called SPICE (Spacecraft, Planet, Instrument, C-matrix (pointing), and Events) to assist which is focused on solar system geometry and related information. The SPICE system includes a large suite of software, mostly in the form of subroutines, used by customers to read SPICE files (also called *kernels*) and to compute derived observation geometry, such as altitude, latitude/longitude, and lighting angles.

SPICE is an essential component in the development and operation of the majority of NASA's solar system exploration missions. It is also being used as an

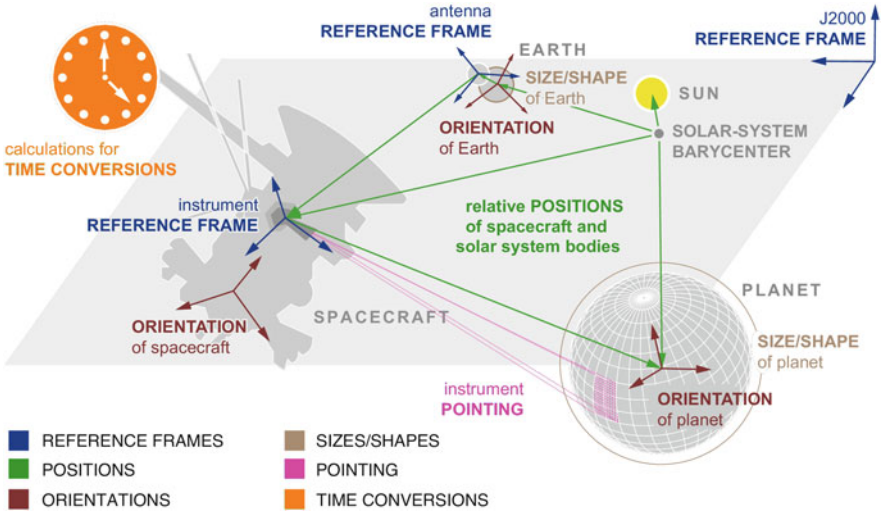

Fig. 4.4 Spice framework. During the capture of an image from a spacecraft, to gather accurate SPICE, the vehicle's ephemeris (trajectory), the planet's or body's ephemeris, orientation and size, the instrument's field-of-view, shape, orientation and lastly the internal timing is critical to understand an image's approximate location on that body. Source: changed after NAIF

adjunct to primary capabilities on some non-U.S. missions such as the ESA Mars Express and Rosetta as well as the Japan Aerospace Exploration Agency (JAXA) Hayabusa and SELenological and ENgineering Explorer (SELENE) missions. Unfortunately, the predicted (as opposed to reconstructed) position and pointing data that are sometimes used to make uncontrolled mosaics can be erroneous, especially for pre-1990s missions, leading to maps with visible mismatches at image boundaries and systematic position errors of many kilometers.

Both ISIS and VICAR are, therefore, used with associated photogrammetric software that estimates improved position, pointing, and planetary parameters based on measurements of matching ground features where images overlap. This process of geodetic control reduces positional errors but cannot, of course, eliminate them entirely. In fact, the lack of ground control for most planets (laser altimetry data now provides an effective substitute for Mars) leads to generally greater errors than are encountered in mapping the Earth. An initial lack of topographic data for each planet or satellite introduces further positional errors in the early maps, because parallax distortions cannot be corrected for. This is no longer an issue for several bodies, for example Mars and the Moon, where images are now routinely orthorectified by using global altimetry datasets.

4.3 Map Projections

Map projections are mathematical equations for mapping a three-dimensional body onto a two dimensional a plane or Cartesian coordinate system. As stated by John P. Snyder , there is no one *best* map projection for mapping and care must be taken choosing a projection that is suitable for the area of study or use given that every projection incurs some type of spatial distortion. The majority of printed maps of the planets and satellites have been based on conformal projections: *Mercator* (Fig. 4.5) for low latitudes, *Polar Stereographic* for high latitudes, *Lambert Conformal Conic*

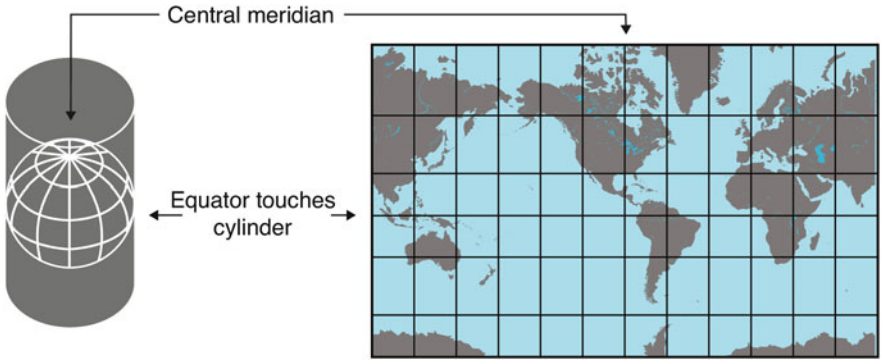

Fig. 4.5 Example showing a mapping from a spherical body to cylindrical map projection. Source: USGS

for intermediate latitudes at small scales, and *Transverse Mercator* for large-scale maps.

Digital map products, on the other hand, have been based on a different set of projections. It is useful to think of these as *database* projections, because their most important properties have been their suitability for holding global datasets. The most important of these qualities are (a) global applicability (i.e., no unmappable areas like polar regions in the Mercator projection); (b) simple formulation; and (c) at least roughly equal area (to avoid oversampling some areas and inflating the volume of the digital dataset). Starting in the 1990s and for many years the preferred database projection for digital map products was the *Sinusoidal projection*. The spherical version of the projection was used in place of the much more complicated and iterative ellipsoidal version of the map projection. However, the Sinusoidal projection was never considered entirely satisfactory for database use because the meridians (apart from the central one) curve relative to the pixel grid. Thus Sinusoidal datasets must be resampled in order for pieces to be combined even at their original scale, and extra pixels must be provided to prevent the formation of gaps along the curved edges of the archived map files when such resampling is performed.

An alternative to Sinusoidal is the Simple Cylindrical projection (in which the grid is simply an equally spaced raster in latitude-longitude coordinates) that eliminates the problems caused by curved meridians. But this greatly oversamples the polar regions, inflating the size of a global dataset by as much as 57%. More recently a useful compromise that has been adopted for an increasing number of digital databases is the Equirectangular projection. This is a generalization of Simple Cylindrical in which latitude and longitude form the grid, rather than having equal latitude and longitude dimensions. The grid cells have dimensions that give equal kilometer lengths at some specified latitude. This is called the *center latitude* though *reference latitude* would be more precise, since it need not be located at the center of a given map. Whereas, in the Sinusoidal projection, the east–west grid spacing in degrees is adjusted on every row to keep the kilometer spacing nearly constant. With the Equirectangular projection, this spacing is constant for a given map file but can be adjusted for files in different latitude zones. The result is a compromise between Sinusoidal and Simple Cylindrical in the complexity and areal distortion of the dataset. Center latitudes may be assigned freely to every latitude zone of the map series (see following section), or they may be chosen to give simple integer relations between the sample intervals in latitude and longitude. In the latter case, center latitude 0° could be used everywhere from 0° to 60° latitude, then 60° center latitude (giving a longitudinal grid spacing twice as large as that in latitude) from 60° to roughly 70°, center latitude 70.5° (giving 3:1 grid spacing) thereafter, and so on. As with the Sinusoidal projection, planetographic and planetocentric variants of the Equirectangular and Simple Cylindrical projections must be distinguished, according to which type of latitude is equally sampled by the rows of the grid.

A relatively recent development has been the acceptance of using the non-conformal, *database* projections for some large-scale maps. This was first done for Venus: the 5° Sinusoidal tiles of the USGS Full-Resolution Radar Mosaics (FMAP) dataset were printed at 1:1,500,000 scale without reprojection on the grounds that

(a) the effort needed to transform the files to conformal projections would have been prohibitive, and (b) the distortions incurred were acceptably small, on the order of 4% at most. This decision highlights the pragmatic character of planetary cartographers. The precedent set by the Venus 1:1,500,000 series was later followed by the Mars Express HRSC team, which adopted the Sinusoidal projection for both digital products and printed maps at scales of 1:200,000 and larger (2° quadrangle size, <1.7% distortion).

One of the criteria for selecting a *database* projection is that it can be used for an entire global dataset. The polar sections of such datasets are entirely adequate as sources of data for resampling to other projections, but are too severely distorted to be directly useful as maps. Separate files are therefore usually provided, showing the polar regions in a more appropriate projection such as Polar Stereographic.

The design of map series is more of an art than a science, and involves reconciling several interlocking sets of constraints. The resolution of available data and the desired density of pixels in the output image are the main factors dictating the choice of a hardcopy map scale (expressed as a ratio of sizes 1:X). Hardcopy scale and the size limits for printing dictate the physical size of map quadrangles for a given body or group of bodies of similar radius. These considerations have led to the definition of a rather large but finite number of quadrangle schemes for mapping the planets and satellites.

In some cases, similar but slightly different schemes have arisen over time, such as the 144-quadrangle scheme for lunar maps at 1:000,000 scale and the 140-quadrangle series of Mars maps at 1:2,000,000 scale. In many cases, schemes for larger scale maps have been developed by subdividing the quadrangles of a smaller-scale map series. Division of quadrangles into quarters is most common, but other schemes have been used, and the division of the near-polar quadrangles is often more complex. Because the set of *round* scales is based on powers of 5 as well as powers of 2, the process of cutting quadrangles into quarters cannot be continued indefinitely. Thus, the 1964-quadrangle scheme for Mars maps at 1:500,000 scale is unrelated to the quadrangles used at larger scales, and it is not further subdivided.

4.4 Nomenclature

Planetary nomenclature is used to uniquely name a feature on the surface of a planet or satellite so that the feature can be easily located and described. The IAU has been the authority of planetary and satellite nomenclature since its first organizational meeting in Brussels, Belgium in 1919. The first goals were to normalize various systems used in lunar and Martian nomenclatures across different countries. The IAU appointed Mary Blagg and several other astronomers to the newly commissioned nomenclature committee, chaired by Herbert Turner, in 1922. The report of this committee, *Named Lunar Formations* was the first systematic listing of lunar nomenclature. Following this first listing several more publications including catalogs for Mars followed. In the 1960s, as detailed remotely sensed

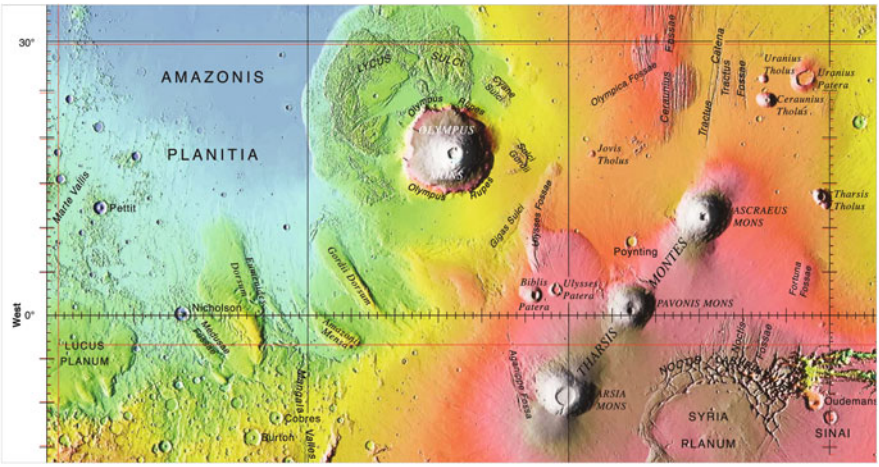

Fig. 4.6 Extracted portion of the original map, at 1:25,000,000 scale, showing nomenclature for only larger features on Mars overlain on a Mars Orbiter Laser Topographic (MOLA) base map. Source: MGS MOLA, NASA/USGS

images became available the need of a single planetary gazetteer of official names became critical for effective scientific communication. The names used within the gazetteer come from cultures all over the world and generally follow a theme for a feature type and body. For example, on Mars, larger craters are named after deceased scientists, especially those who have contributed significantly to the study of Mars as well as writers and others who have contributed to Mars' lore (Fig. 4.6). For Mercury, large valleys are named after abandoned cities (or towns or settlements) of antiquity. The current nomenclature database is managed by the *Astrogeology Science Center* on behalf of the IAU.[3] It is recommended that maps created for digital or hard-copy distribution should display nomenclature completely and accurately when possible, as defined by the map's scale.

4.5 Digital Mapping

Both ISIS and VICAR can be used for overlay analysis and geologic mapping in addition to their primary functions of producing digital map bases, but in recent times such functions have mainly been handed over to more capable *Geographic Information Systems (GIS)* applications. GIS are computerized systems for the storage, retrieval, manipulation, analysis, and display of geographically referenced data that can be tied to a spatial location, including tabular or other information associated with that location. However, for GIS-applications to be primarily useful

[3] http://planetarynames.wr.usgs.gov.

for research, the data should be radiometrically and photometrically calibrated and spatially registered to the body as explained above (Fig. 4.3). But beyond those initial data preparation steps, a GIS should be able to handle nearly all digital mapping projects including data ingestion, storage, creation, manipulation, multi-layered analysis, visualization, and data output (whether it is digital or hard-copy), through an easy to use interface.

GIS applications have been used for planetary projects for many years. Planetary researchers now rely on GIS technologies because it offers the ability to overlay multiple datasets and to help facilitate characterizing and analyzing a planet's geologic and structural history. The utility of GIS has expanded to include on-line data dissemination, mission planning, mission support, and modeling of planetary processes. GIS analysis lets the user ask: where, why, and how: for example, where are distinct features coincident, why do these structures exist at this latitude range, or how does this process affect another?

In recent years, there has been rapid growth for GIS technologies and spatial analysis in extraterrestrial research. Increasingly, outside the realm of planetary geologic mapping, specialists such as astrobiologists, planetary protection experts, and planetary hardware designers are striving for more autonomous robots that can catalog and analyse multiple types of information without human interaction, also using GIS technology.

4.5.1 Geologic Mapping

Geologic maps present the fundamental syntheses of interpretations of the materials, landforms, structures, and processes that characterize planetary surfaces. They also provide a regional or global contextual framework for summarizing and evaluating thematic research. For robotic and human planetary exploration, geologic maps are, and will continue to be used for specialized investigations such as targeting regions of interest for data collection and for characterizing sites for landed missions.

Planetary geologic maps have been published by the USGS starting with Hackman's 1962 geologic map of the Kepler region of the Moon. Since then, geologic maps of many planetary bodies have been prepared at a variety of scales and projections using the most continuous and highest-resolution image and topographic bases (Fig. 4.7).

As with terrestrial geologic maps, all planetary geologic maps published by the USGS now are primarily digital products using GIS software and file formats. GIS mapping tools permit easy spatial comparison, manipulation, and analysis of multiple raster image, gridded, and feature (vector) data sets. In addition, GIS software also permits the development of project-specific tools and the sharing of geospatial products among researchers and the general public. Examples of GIS-specific geologic mapping and analysis tools include geodesic calculators that yield accurate areas, lengths, and angles regardless of the map projection and specialized digitizing tools to identify and analyze the density of impact craters on particular surfaces.

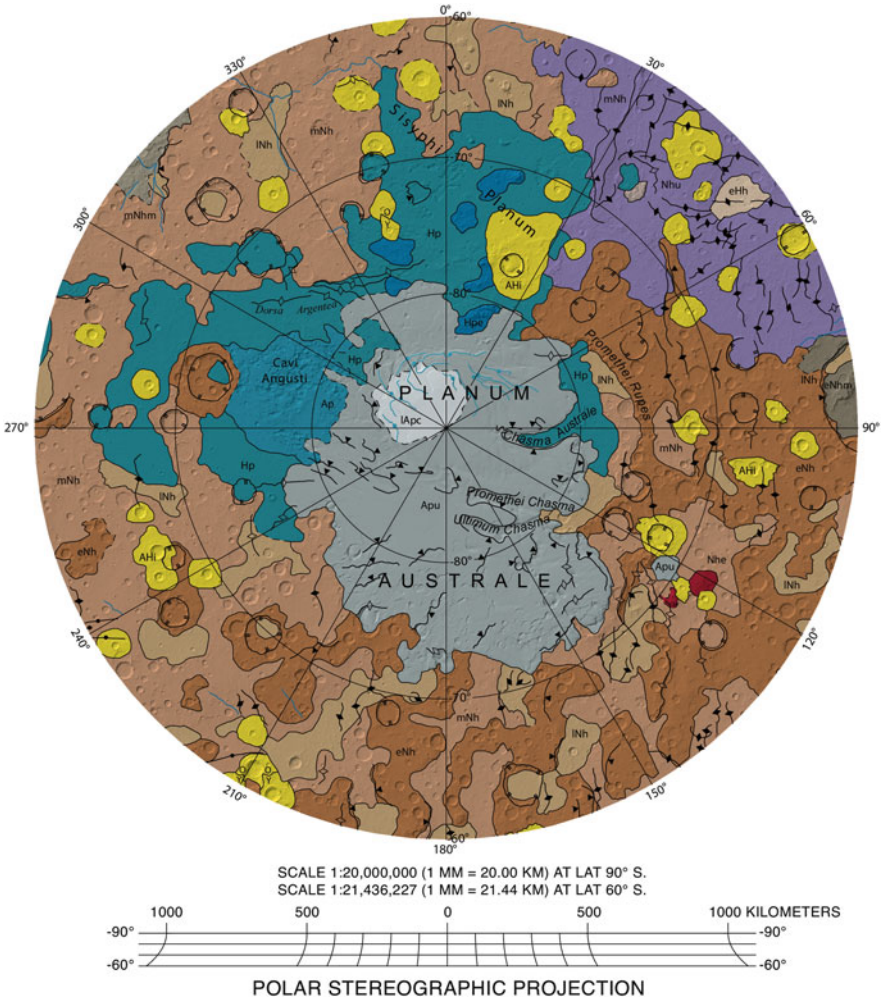

Fig. 4.7 Extracted portion showing the South Pole of the original global Geologic Map of Mars. Vertex spacing of drafted line work was set at 5 km (4 vertices per millimeter at 1:20,000,000 scale) and the minimum feature length accepted was 100 km. Source: Tanaka et al. (2014)

Though the process of modern geologic mapping is generally well-known, the detailed procedural components of producing high-quality standardized, peer-reviewed, and technically edited geologic maps are complex and involve a wide range of data, cartographic software tools, technical procedures and publication requirements. Consistent adherence to community standards regarding mapping methods and representation is critical and must be thoroughly considered, reported on, and reviewed. These methods apply to with respect to data collection, attribution, symbology, documentation, and distribution formats partially discussed

REF NO	DESCRIPTION	SYMBOL	CARTOGRAPHIC SPECIFICATIONS*
25.94	Raised rim of larger impact crater, planetary—Hachures point into crater		all lineweights .3 mm hachure height .75 mm; spacing of hachure pairs .5 mm .375 mm
25.95	Raised rim of smaller impact crater, planetary		lineweight .3 mm
25.96	Raised rim of impact crater, planetary—Showing visible ejecta blanket		lineweight .15 mm
25.97	Degraded impact crater rim, planetary (1st option)		lineweight .3 mm dash length 1.0 mm; spacing .5 mm
25.98	Rimless impact crater, subdued impact crater rim, degraded impact crater rim (2nd option), or buried impact crater rim, planetary		lineweight .3 mm long dash 4.0 mm; short dashes .2 mm; spacing .5 mm
25.99	Secondary impact crater chain and cluster, planetary		lineweight .25 mm dash length 1.5 mm; spacing .5 mm
25.100	Basin ring, planetary		lineweight .375 mm dash length .75 mm; spacing .75 mm
25.101	Central peak of impact crater, planetary (1st option)		ellipse width 1.875 mm; height 2.625 mm 1.5 mm all lineweights .2 mm
25.102	Central peak of impact crater, planetary (2nd option)	+	2.375 mm all lineweights .2 mm

Fig. 4.8 Partial page extracted showing only a few standardize planetary symbol as defined in the FGDC Digital Cartographic Standard for Geologic Map Symbolization. Source: FGDC

next. Using these standardized methods allows these products to be readily understood and used across the community.

4.5.2 Attribute and Symbology Standards

As mentioned previously, the IAU defines the recommended rotation rate, spin axis, prime meridian, and reference surface for planets and satellites; however, their oversight does not cover other standards essential for digital mapping including common feature attributions, feature symbols, recommended mapping scales and finally metadata or the documentation of the data. When possible, it is recommended that digital maps use these standards so that maps are more easily understood when prepared by different authors and across different products.

Feature attributes and their assigned symbols (Fig. 4.8) for planetary digital maps are commonly defined in the *Digital Cartographic Standard for Geologic Map Symbolisation* prepared by the USGS for the *Federal Geographic Data Committee (FGDC)*[4] For example, recommended attributes for geologic contacts or geologic unit boundaries include attributes like certain contact, approximate contact, inferred contact, or gradational contact (see Chap. 2). These geologic contact types are then mapped to unique line representations like a solid black line or, for approximate, a

[4]http://www.fgdc.gov.

dashed black line. Each feature attribute can be visualized by these different lines, point symbols, solid or hachured fill patterns for polygons as either stand-alone symbols or possibly a combination of symbols. Fortunately for planetary mappers, these symbols are primarily based on the same set of attributes and symbols as used for Earth. This ultimately facilitates the understanding of geologic or thematic planetary maps because readers are familiar with the names and symbols.

4.5.3 Mapping Scale and Data Collection

While automatic techniques to facilitate and collect surficial features with geologic mapping are maturing, most maps are still created by manually drawing the features or geologic contacts using computer draw-screen, tablet or input devices like computer mice (Fig. 4.7). Several technical issues arise with digital maps that were previously not issues when using pen and paper. Mappers should be cognizant of the actual data collection rate while mapping. This includes the minimum feature size collected as well as the data collection rate while drawing lines or polygons. Features too small will clutter the map with unnecessary details and too many vertices along the line will generally bog down the digital map. If the collection rate is not great enough, the map will not show adequate detail.

Recommended line-work should have a vertex spacing of ~0.3 mm at map scale (equivalent to 300 m for a 1:1,000,000-scale map). Maps should also be collected at a consistent scale so that the collection rate is similar across the map. The recommended mapping scale is usually a factor of 2–5 larger than the published map to ensure adequate precision (e.g., digitally map at 1:200,000–1:500,000 for a 1:1,000,000 published map). If necessary, GIS tools can be applied to existing line-work to generalize (remove redundancy) and to smooth line-work to achieve the desired result, such as rounded corners. Also, single features, like craters, should generally be at least 2 mm wide at map scale. Reasonably sized cutoffs should also be defined for line feature lengths (for example, 1 or 2 cm long at map scale). Points can be used to show the distribution of important features such as craters and shields that are too small to map.

4.5.4 Metadata

Metadata is the necessary ancillary documentation that describes each GIS layer in a map, including rationale, authorship, attribute descriptions, spatial reference, and other pertinent information as required by the metadata standard. The *FGDC* in cooperation with the *International Organization for Standardization (ISO)* endorses the development, use, sharing, and distribution of geospatial data. When this standard is applied correctly, users should have a significantly clearer understanding of the described data and it should be more easily applied to their research goals. The

objective for any metadata standard should be to provide organisation of electronic resources, resource discovery, interoperability, digital identification, and archive and preservation.

During 1992, the *FGDC* held a forum on the many aspects of metadata, including efforts to standardize metadata, the uses of metadata, and systems used to provide metadata. One of the outcomes of this forum was an offer from the *American Society for Testing and Materials (ASTM)* to develop a draft standard. *ASTM*, like *ISO*, is globally recognized for the development and delivery of international standards. Once finalized in 1994, the newly created *Content Standard for Digital Geospatial Metadata (CSDGM)* quickly became the official metadata standard for geospatial data in the United States. The major sections in the *CSDGM* standard include the following:

- Identification—basic information about the data set
- Spatial Data Organization—how the data is organized
- Spatial Reference—description of the Cartesian system
- Entity and Attribute—content of the data including entity types and attributes
- Distribution—the distributor of and options for obtaining the data
- Metadata Reference—currentness of the metadata and responsible party

In 1999 the *ISO Technical Committee (TC) 211 Geographic Information/ Geomatics* was tasked with harmonizing the *CSDGM* with other geospatial metadata standards and a range of *de facto* standards that had emerged to address new requirements for geospatial documentation. The result was the publication of *ISO 19115: Geographic information—Metadata*. The *ISO 19115* standard goes further than the *CSDGM* because it provides an international method of documenting geospatial data products, includes international references (language and characters), addresses new data structures, and supports elements for related data applications and services. Unfortunately, the *ISO 19115* is an abstract standard which specifies the definition and relationship of the elements but does not define how the content is organized into a formal record. Without this concrete structure for a formal record, the standard cannot be endorsed for implementation. Thus another *ISO Technical Committee TC 211* was tasked to generate an Extensible Markup Language (XML) implementation schema for the *ISO 19115* standard. This follow-on standard released in 2003, called *ISO 19139—The Geographic information—Metadata—XML schema implementation*, specifies the metadata structure in XML which allows one to validate and more easily exchange *ISO 19115* geospatial metadata.

4.5.5 Analysis

GIS excels at overlaying multiple data sets which can highlight relationships that were not otherwise obvious. The idea that data creates more data is especially true when using a GIS. By using standard GIS tools, existing data can be used alone

or intersected with other datasets to generate new layers. For example, topographic layers can be used to generate slope and aspect maps, contours, hillshades, watersheds, drainage delineations, or viewshed layers (line-of-sight analysis). Datasets, when combined together, can generate almost unlimited combinations of new layers. For example, combining a slope map with a thermal inertia map can help generate a landing site hazard map where slopes too steep or areas too dusty should be avoided. Intersecting the number of craters with geologic unit types can help to define the age of that that unit. Examples of this type of analysis are shown in Chaps. 8 or 11.

Take-Home Messages

No serious extraterrestrial geologic study can be accomplished in the absence of cartography.

Cartographic standards such as uniformity of coordinate systems, accurate horizontal and vertical positioning, and mapping methods, scales and schema are vital components to allow data usage across many facilities including surface exploration by robots and humans.

There has been rapid growth and continual evolution for mapping technologies and spatial analysis in extraterrestrial research.

For mapping applications to be useful for research, the data should be radiometrically calibrated and spatially registered and controlled to the body.

Suggested Readings

Acton, C.H.: Ancillary data services of NASA's navigation and ancillary information facility. Planet. Space Sci. **44**(1), 65–70 (1996). doi:10.1016/0032-0633(95)00107-7

Archinal, B.A., A'Hearn, M.F., Bowell, E., Conrad, A., Consolmagno, G.J., Courtin, R., Fukushima, T., Hestroffer, D., Hilton, J.L., Krasinsky, G.A., Neumann, G., Oberst, J., Seidelmann, P.K., Stooke, P., Tholen, D.J., Thomas, P.C., Williams, I.P.: Report of the IAU working group on cartographic coordinates and rotational elements: 2009. Celest. Mech. Dyn. Astron. **109**(2) (2011). doi:10.1007/s10569-010-9320-4

Greeley, R., Batson, R.A.: Planetary Mapping. Cambridge Planetary Science Series, vol. 6, 312 pp. Cambridge University Press, Cambridge (1990)

Hare, T.M., Rossi, A.P., Frigeri, A., Marmo, C.: Interoperability in planetary research for geospatial data analysis. Planet. Space Sci. (2017). doi:10.1016/j.pss.2017.04.004

Hargitai, H.I., Shingareva, K.B.: Planetary nomenclature: a representation of human culture and alien landscapes. In: Advances in Cartography and GIS science. Lecture Notes in Geoinformation and Cartography, vol.6(4), pp. 275–288. Springer, Berlin (2011). doi:10.1007/978-3-642-19214-2_18

McMahon, S.K.: Overview of the planetary data system. Planet. Space Sci. **44**(1), 3–12 (1994). doi:10.1016/0032-0633(95)00101-8

Naß, A., Di, K., Elgner, S., van Gasselt, S., Hare, T., Hargitai, H., Karachevtseva, I., Kersten, E., Manaud, N., Roatsch, T., Rossi, A.P., Skinner Jr., J., Wählisch, M.: Planetary Cartography and Mapping: Where We Are Today, and Where We Like/Have To Go? International Architecture Photogrammetry. Remote Sensing Spatial Infrastructure Science, vol. XLII-3/W1, pp. 105–112 (2017)

Shevchenko, V., Rodionova, Z., Michael, G.: Lunar and Planetary cartography in Russia. Astrophysics and Space Science Library, xiii+145 pp. Springer, Berlin (2016). doi:10.1007/978-3-319-21039-1

Shoemaker, E.M., Hackman, R.J.: Stratigraphic basis for a lunar time scale. In: Kopal, Z., Milchailov, Z. (eds.) The Moon, pp. 289–300. Academic, London (1962)

Snyder, J.: Map projections–a working manual. In: U.S. Geological Survey professional paper, vol. 1395, U.S. Geological Survey (1987). https://pubs.er.usgs.gov/publication/pp1395

Tanaka, K., Skinner, J., Dohm, J., Irwin, R., Kolb, E., Fortezzo, C., Platz, T., Michael, G., Hare, T.: Geologic map of Mars, scale 1:20,000,000. U.S. Geological Survey Scientific Investigations Map, vol. 3292, 43 pp. U.S. Geological Survey (2014). doi:10.3133/sim3292

Wilhelms, D.: Geologic history of the moon. U.S. Geological Survey Professional Paper, U.S. Geological Survey (1987). http://pubs.er.usgs.gov/publication/pp1348

Chapter 5
Ground Truth

Angelo Pio Rossi and Stephan van Gasselt

5.1 Introduction

The robotic exploration of the Solar System, similarly to remote-sensing observation and monitoring of our own planet, is key to capture the broad, large-scale perspective and reconstruct the evolution of bodies located millions or even billions kilometres away from the Earth. Such distances in space and time separating us from the geological objects of study are challenging. Interpretation of remotely-sensed data can be affected by large errors, like any measurement, but most importantly by interpretation bias.

The *present is key to the past* (Chap. 2) and the Earth is our key to interplanetary uniformitarianism, but *ground truth*, on Earth and even more on other planets and moons is the tool of choice to develop and understanding of planetary environments and processes. Fieldwork, in particular if robotic, does not necessarily provide final answers. Typically, extended fieldwork complementing remote sensing and well-contextualised sample return form the best combination for planetary geological exploration. Such combination, however, is rare, and has thus far only been achieved for the Moon while Mars might be the next candidate.

The terrestrial analogue field perspective (Chap. 2) is important, both for scientific and operational aspects, although there are limits in both cases. Conditions might be relatively close in terms of temperatures for certain analogues, but very far with respect to other parameters, such as gravity, or surface pressure.

A.P. Rossi (✉)
Jacobs University Bremen, Campus Ring 1, 29795 Bremen, Germany
e-mail: an.rossi@jacobs-university.de

S. van Gasselt
National Chengchi University, No. 64, Sec 2, ZhiNan Rd., Wenshan District, Taipei 11605, Taiwan
e-mail: svg@nccu.edu.tw

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_5

In the case of the terrestrial planets and the Moon similarities might be relatively close, e.g. lithologically, when one compares Venusian plains to terrestrial basalts. For small bodies such as asteroids and comets (Chap. 13) it is much more challenging to simulate surface (or subsurface) conditions on the Earth. Complications are further increased in the case of icy satellites hosting subsurface oceans (Chap. 12), where analogies, operationally and also geologically, are conceivable.

In all the cases outlined above remote sensing observations tend to be ambiguous, thus, the need for independent *ground truth* arises. Sample return from extraterrestrial objects (Chap. 3) is one of those cases where such field control, with in-situ measurements are of paramount importance. Going into the field is central in Earth geology (Chap. 2), but in planetary exploration it is mostly achieved with robotic platforms (Chap. 3) except for the thus far unique case of the Moon. Benefits of in-situ verification and sample acquisition for further lab analyses, impossible, difficult or overly expensive to be performed otherwise, are several: one can design and deploy exploration platforms lighter and relatively less complex and easier to operate. Autonomous robotic exploration for geological purposes, however, does not exist yet, and the only 'platform' able to do that is human, but in the forthcoming decades this might change.

Until today human-only geological exploration (Fig. 5.1a, b) has only been achieved on the Earth, and even terrestrial exploration begins to rely on robotic resources, such as drones. In addition, tools such as hand-held LIDAR and spectrometry of various kind support basic geological mapping and goes far beyond the classical use of topographic maps as a basis for surveying (Chaps. 2, 4).

The NASA Apollo Program constitutes to date—half a century from its beginning—an unparalleled coupling of effort and achievement that shaped Planetary Geology (Chap. 1) and sparkled the development of many disciplines within. Astronauts, either originally or specifically trained in the geosciences (Fig. 5.1c) allowed for the collection of data and samples of high-quality, with appropriate geological context. Its heritage is still being geologically exploited, its samples are still used for making new discoveries and any future planetary exploration, on the Moon itself, on asteroids or on Mars, will have to measure itself with the scale of success of the Apollo program.

5.2 Lander and Rover Exploration

Landers and mobile rovers are the most well-known and common platforms to collect field data on planets other than the Earth. They look back on more than half a century of developments and field experience, following—and often leading—the developments of robotics and autonomy, as well as the range of experiments available to characterise extraterrestrial surfaces (Fig. 5.1d).

The most direct and somehow violent way to perform ground truth is to trigger a surface or subsurface effect remotely with impactors which are potentially equipped with sensors. This has been done both in the early phases (also recently with

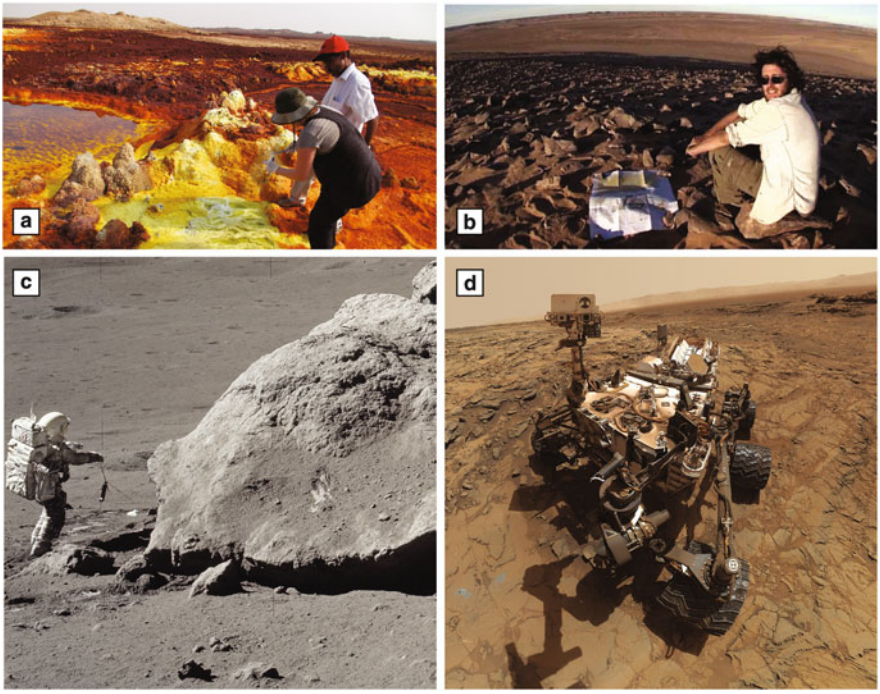

Fig. 5.1 Human vs. robotic ground truth: The level and amount of equipment is variable, from just hammer or simple sampling device to specialised tools and analytical facilities. The degree of autonomy is also highly variable: (**a**) B. Cavalazzi sampling hydrothermal hot springs in the Dallol crater (Afar, Ethiopia). (**b**) A.P. Rossi performing panoramic observations over the inner rings of the Richat structure (Mauritania). (**c**) Apollo 17 observations in the vicinity of a large ejected block. (**d**) MSL rover self-portrait mosaic of MastCam images: The dust cover is comparable with that shown in (**b**). Source: (**a**) B. Cavalazzi. (**b**) R. Sabbadini. (**c**) NASA Apollo photo as17-140-21496. (**d**) ASA Photojournal image PIA19920

LRCROSS) of Moon and cometary exploration (Deep Impact). Artificial impacts can expose subsurface material, as well as provide information on the target's material nature, composition and strength (Chap. 7). Some impacts or atmospheric plunges have been planned as end of a mission lifetime (SMART-1 on the Moon, Messenger on Mercury, Galileo on Jupiter and soon Cassini on Saturn), but the observations of their effects has been more problematic, since they were the only platform orbiting their target body, and thus not able to observe themselves while impacting the surface. A somewhat similar, low-speed end-of mission was planned and performed for ESA Rosetta. Specific cases of spacecraft impactors deliberately deployed and observed by the mother spacecraft were used e.g. with NASA Deep Impact.

The non-destructive deployment of static landers equipped with scientific experiments on the surface of a Solar System body is much more complex. The first body where this was achieved is the Moon for obvious reasons. After several ill-fated

attempts in the late 1950s to get to the Moon with impactor and orbiter probes, the very first robotic impactor landing on another planetary object was the Soviet Luna 2 in 1959. Its science return was relatively limited but the pressure was raised on the US, in particular after the 1957 Sputnik event. A month later, the Soviet Luna 3 spacecraft was the first to send images of the lunar farside back to Earth. It was not until 1964 that the US successfully employed their Ranger program, 3 years after the Apollo program had been officially announced. The Ranger impactors, equipped with a set of vidicon cameras operating with different focal lengths, were also designed as impactor missions and targeted at the Moon. Despite malfunctions, Ranger 4 hit the farside surface but only the last 3 Ranger missions successfully impacted and returned sequences of descent images to Earth in late 1964 and early 1965. It was 1966 when Luna 9 made the very first soft touch down and returned first images from the surface of the Moon. Just a few months later, the US successfully targeted the seven Surveyor landers to the Moon which all succeeded and which paved the way for the successful Apollo program. During the six Apollo landings between 1969 and 1972, the Soviet Union sent several spacecraft for autonomous operation and sample return to the Moon. Finally, Luna 16, 20 and 24 succeeded to bring back rock samples. The successful Luna 17/Lunokhod 1 and Luna 21/Lunokhod 2 rovers ended that area of lunar exploration in 1973, 1 year after termination of the Apollo program. It took until 2013 when the Moon was visited again by spacecraft. At that time the Chinese Space Agency (CNSA) sent the Chang'e 3 lander mission to the Moon where it deployed the rover Yutu to perform surface operations.

Once a static lander is deployed on the surface and with enough power supply (either battery-, solar- or nuclear-powered), the lifetime of the lander itself and the data volume and length of the data time series collected strongly depend on the amount of energy and its replenishment, the degradation of the hardware and its intrinsic technical life span (e.g. maximum battery recharge cycles) as well as the harshness of the surface environment with respect to the lander life span itself (extreme temperatures, pressures, or alike). Attached to a static lander several types of experiments can be run, optimally (such as in the case of Viking 1, 2 and Phoenix landers) with a robotic arm to either deploy an experiment on the surface or collect surface or subsurface materials to be analysed on the lander itself. Simple tethered rovers can also be attached to a lander, providing energy and eventual communication to a small rover.

Wheeled rovers can range from very small, toy car-sized (e.g. Mars-2 rover with 4.5 kg mass) to large real car-sized objects (e.g. MSL Curiosity with 900 kg, and Mars 2020). Their capabilities, range and operational life are typically directly proportional to their size. Operations ranging from hours to years (even over a decade as in the case of MER) are possible. Mobility opens a much wider range of possibilities for observing diverse geology and terrain types. With increasing mass and volume available for experiments, the number, quality and possible complexity of ground truth-performing instruments increases. Sample acquisition can be more advanced (e.g. surface coating removal, shallow or deeper subsurface sample acquisition), allowing also sample preparation and processing complex

enough to run small-scale laboratory procedures and analyses inside the chassis of the rover (e.g. X-Ray diffractometry, mass-spectrometry).

The accuracy level or type of measurements are nevertheless limited on a rover, even if as big as a small bus. The following step in terms of capability, where rover-based experiments cannot deliver, is to select using the experiments on board the rover itself, collect and cache samples for eventual return to Earth. In this case, ground truth literally brings samples back to ground. Robotic caching and return of sample from the Moon has been achieved decades ago (USSR LUNA), human-robotic one as well (so far the most effective both qualitatively and quantitatively). Mars is far more difficult in this respect. Further and farther than Mars, the effort is increasingly complicated.

Almost regardless the type of ground truth experiments and sample processing, analysis or caching capabilities, what largely makes the difference is how much diversity of geological units and materials can be accessed and how good it can be characterised. Beyond static landers and wheeled rovers, several other landing exploration options have been developed, particulary for low-gravity environments such as those on small bodies: hopping has been developed for JAXA Hayabusa and Hayabusa-2 missions, e.g. with the DLR MASCOT hopping lander. Further ahead, different types of locomotion are envisaged, e.g. for low-gravity bodies more or less rugged terrains hopping and rolling or, for higher-gravity targets, such as planets, legged robots: their terrestrial applications are relatively new and their planetary counterparts not common yet.

More recently, flying drones or UAVs have received considerable attention. They can be considered as semi-autonomous to autonomous robotic platforms complementing ground truth operations and have been rapidly spreading in Earth research. They are well-suited for planets and planetary satellites with an atmosphere such as Venus, Mars, and Titan.

The nature of ice and volatile-rich Outer Solar System bodies (Chap. 12) requires a special design of landers, also due to planetary protection aspects of bodies such as Icy Satellites (Chap. 14). For most of the categories above, the first and closest testing ground is the closest extraterrestrial object: the Moon.

Testing ground truth platforms and related experiments for low gravity environments (asteroids, comets) can be performed in condition of microgravity, e.g. using parabolic flights or drop tests). The choice of landing sites (Chap. 3) for ground truth is extremely important: although any field site on a previously unexplored planet or planetary satellite typically triggers a leap in geological understanding, the cost and effort involved in landing missions require a careful and well-designed selection process. When very little is known about the surface, more flexibility is needed: for example the Huygens probe was not known to land on solid or liquid (methane) surfaces on Titan, thus, it was designed to be able to float as well—a feature that was not needed in the end. The information available in order to select a meaningful landing site can vary substantially: from very limited data such as for the first set of lunar missions, to extensive multi-mission, multi-experiment datasets and models, such as in the case of later and near-future Mars rover missions.

The number of observations needed during ground truth for characterising aspects of the local geology (morphology structure, composition, age, state of activity etc.) is a key aspect: if the final aim is to collect only few tens of samples the number of ground-based remote sensing observations will be very high (10^2–10^3), the number of contact measurement in-situ is on the order of 10^2, as based on the MER Spirit experience: such order of magnitude is respected on similarly equipped missions, with comparable conditions (e.g. MSL).

5.3 The Moon

The Moon is the Solar System body with the overall best level of ground truth, performed through robotic and human in-situ exploration as well as sample return, carried both automatic probes and by astronauts. The number of landing sites on the Moon is higher than any other Solar System body (Fig. 5.2). It is a body with very limited large-scale geological heterogeneity and a limited number of terrains (anorthositic highlands vs. basaltic lowlands) with substantial regional and local complications due to both impact, volcanic/tectonic processes and solar forcing. Also, the absence of any sizeable atmosphere is both an asset, limiting complications dealing with time-varying atmospheric conditions such as those on Mars, and an issue, not providing any deceleration for eventual incoming probes.

The geological significance of the Moon has multiple facets: it is centered on its capacity to retain large portions of geological history of the early Inner Solar

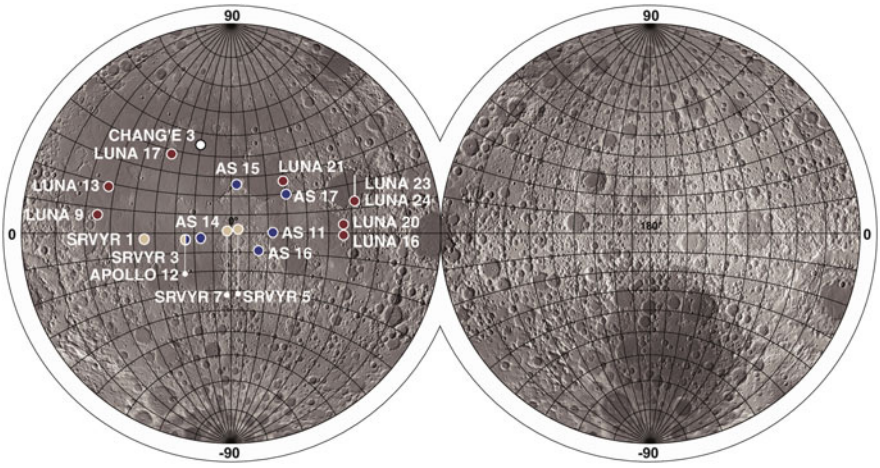

Fig. 5.2 Location of spacecraft landed on the Moon. Topography is color-coded (*dark* = low, *bright* = high) shaded relief in stereographic projection with central latitude and longitude at 0°N/0°E and 0°N/180°E. *Symbol colors* refer to mission program. Source: lander positions: NASA NSSDC, Wikipedia, on NASA LRO/LOLA hillshade

System that is either lost or unaccessible elsewhere. The Moon is also in the position to shed light on Earth's unique evolutionary path among all terrestrial planets. The only body in which absolute calibration of crater size-frequency dating has been directly performed is the Moon (Chap. 7). On more strictly exploration terms, the Moon can be the most proximal large-scale extraterrestrial outpost beyond Earth, as well as a possible remote, but still relatively close, source of resources valuable for further exploration (Chap. 15), as well as a possible base for it.

Several, if not most, of those key geological features were not well-known before the robotic exploration of the Moon started. Pre-spacecraft astronomic observations could highlight already large provinces and the features of impact basins and craters, including their inner morphology, ejecta and structural features. Key concepts and hyphotheses in planetary science such as the giant impact hypothesis and models such as the magma ocean developed from Moon-based data and samples, later applied elsewhere in the Solar System.

The scientific importance of the Moon for Earth's history and the Solar System as a whole has been understood thanks to the exploration itself. Moon exploration, both remote-sensing, in-situ and via sample return is what effectively kick-started Planetary Geology as a discipline (Chap. 2). It all began with the efforts culminating in the US Apollo and USSR Luna programs. With a pattern similar to that of Mars' exploration, but more pronounced, the Moon had a peak of ground truth and preparatory remote sensing missions in the 1970s, followed by a long gap until the mid 1990s and a renewed, multi-faceted exploration in recent years and upcoming future.

Only from the lunar nearside lander units can communicate directly with Earth without orbital relay. Consequently, lander missions have been place on the nearside but new approaches to send landers to the farside will follow in the upcoming years (Fig. 5.1).

The types of platforms used for ground truthing the Moon is extensive, from penetrators to static landers, particularly in the early phases, up to robotic rovers, such as USSR's six-wheeled Lunokhods. Ground truth platforms had to withstand lower gravity, no surface pressure and a harsh thermal environment with temperature ranges across illuminated and shadowed areas of about 300°K. The range of in-situ experiments deployed on the Moon is also the largest, although most of them date back to the 1970s, therefore they do not include the relatively recent technological advances available to Mars lander/rover exploration the last decade has brought up. In addition to lander-based remote sensing and photography aimed at characterising surface morphologies, experiments targeted at soil mechanics, chemical composition, mostly robotic, as well as passive and active geophysical experiments (Chap. 3). Most of them were deployed by astronauts and collected long time series of data, until even today. Sample collection and return was performed in both automated and manual ways (Figs. 5.3 and 5.4).

As outlined above, early robotic exploration by the US and USSR included fly-by attempts, orbiters and also employed hard and soft landers in the 1960s. After the initial success by the USSR to reach Earth orbit in 1957, to photograph the lunar farside for the first time, and to impact on the lunar surface in 1959 the

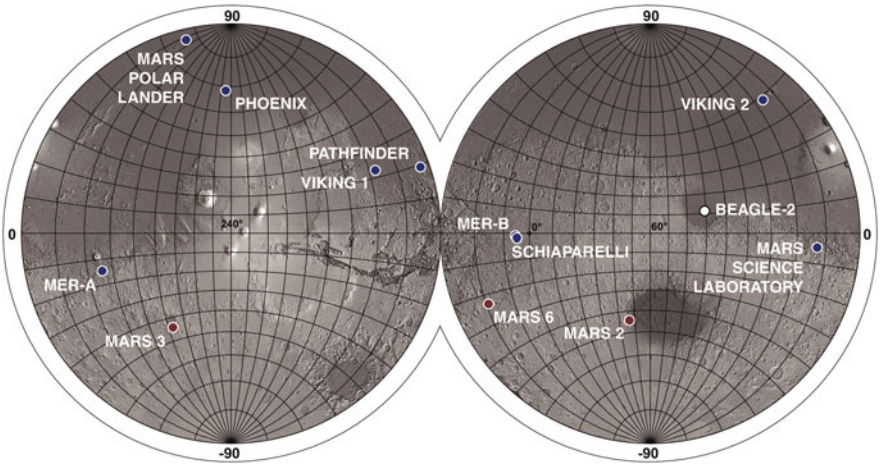

Fig. 5.3 Location of spacecraft landed on Mars. Topography is color-coded (*dark* = low, *bright* = high) shaded relief in stereographic projection with central latitude and longitude at 0°N/240°E and 0°N/60°E. *Symbol colors* refer to mission program. Source: lander positions: NASA NSSDC, Wikipedia, on NASA MGS/MOLA hillshade

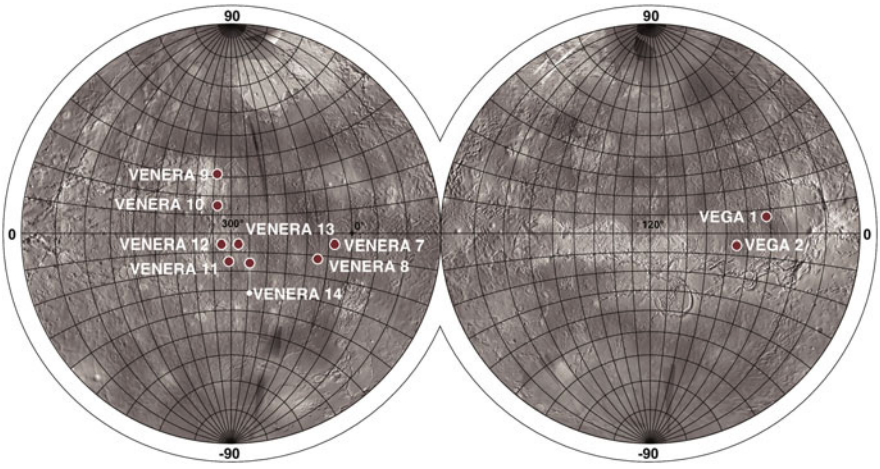

Fig. 5.4 Location of spacecraft landed on Venus. Topography is color-coded (*dark* = low, *bright* = high) shaded relief in stereographic projection with central latitude and longitude at 0°N/300°E and 0°N/120°E. *Symbol colors* refer to mission program. Source: lander positions: NASA NSSDC, Wikipedia, on NASA/Magellan hillshade

US followed with the Ranger and Surveyor program as precursor to the Apollo investigation. They were complemented by the short-lived but overwhelmingly successful Lunar Orbiter program between 1966 and 1977 during which the Moon was imaged at highest resolution. The spacecraft of Surveyor 3, landed in 1967, was revisited by the Apollo 12 astronauts in 1969 after a pinpoint landing in walking

distance away from the landing site. Samples of the Surveyor spacecraft were returned to Earth for inspection and characterization of the lunar environment. The ranges of objectives during the early investigations were broad, largely linked to the characterisation of the lunar surface, imaging and the identification of potential landing sites and their suitability and safety for subsequent human-robotic landings. Besides imaging, the investigation of geophysical and astrophysical properties of the lunar environment were the main focus of Soviet missions during the 1960.

The Apollo program initiated in 1961 and lasting until 1972 was a large, complex and ambitious set of missions that enabled human exploration of multiple locations on the nearside of the Moon. It included both orbital and landing elements, the first of which did not reach the Moon but explored the space environment and tested systems and procedures in Earth orbit (Apollo 7, 9), orbited the Moon and returned to Earth (Apollo 8, 10, 13, the latter was planned to land originally) and 6 successfully orbited, landed and returned (Apollo 11, 12, 14, 15, 16, 17, in Fig. 5.2). Luna, roughly synchronous with Apollo, included several elements from basic penetrators to complex, autonomous sample return missions (Luna 16, 20, 24, in Fig. 5.2).

Experiments carried out or deployed by astronauts during the Apollo missions covered areas beyond geology, such as measuring in-situ the solar wind (SWC), but the majority of instruments and operations were focused on the geology. Equipment included: soil mechanics experiments, chemical and mineralogical experiments for in-situ characterisation of samples to eventually collect, geophysical experiments. Certain experiments were deployed and left to collect data, others were used to support fieldwork and sample collection.

Apollo samples and to a lesser extent in-situ data are still being exploited in these years. The preparations and planning of field activities for Apollo was thoroughly reviewed and procedures tested well in advance during the geologic training of astronauts on analog sites on Earth (Chap. 2).

Exploration was typically conducted in traverses, planned in advance and moderately adjusted based on local boundary conditions and geological variability at field scale. Field equipment included rock and regolith handling tools, such as hammers, chisels, scoops, rakes, as well as corers and drillerscapable of reaching several decimeters to few meters of the subsurface. Aids for photographic documentation was provided by scales and gnomon, recording local illumination conditions. Geophysical equipment used at discrete locations included magnetometers, seismometers and gravimeters, in addition to active seismic transmitting and receiving systems (Fig. 5.5).

A substantial focus on resources was related the renewed exploration of the Moon in the 1990s and later years, with both Clementine and Lunar Prospector missions, which paved the way for the recent multinational orbiters with high-resolution imaging from NASA (LRO), Japan (Kaguya/SELENE), China (Chang'e), as well as impacting elements (LCROSS), focusing on volatiles. All of those constitute a set of precursors, at different stages, of human exploration, thus characterising with increased resolution, globally or locally, the surface features, environment and

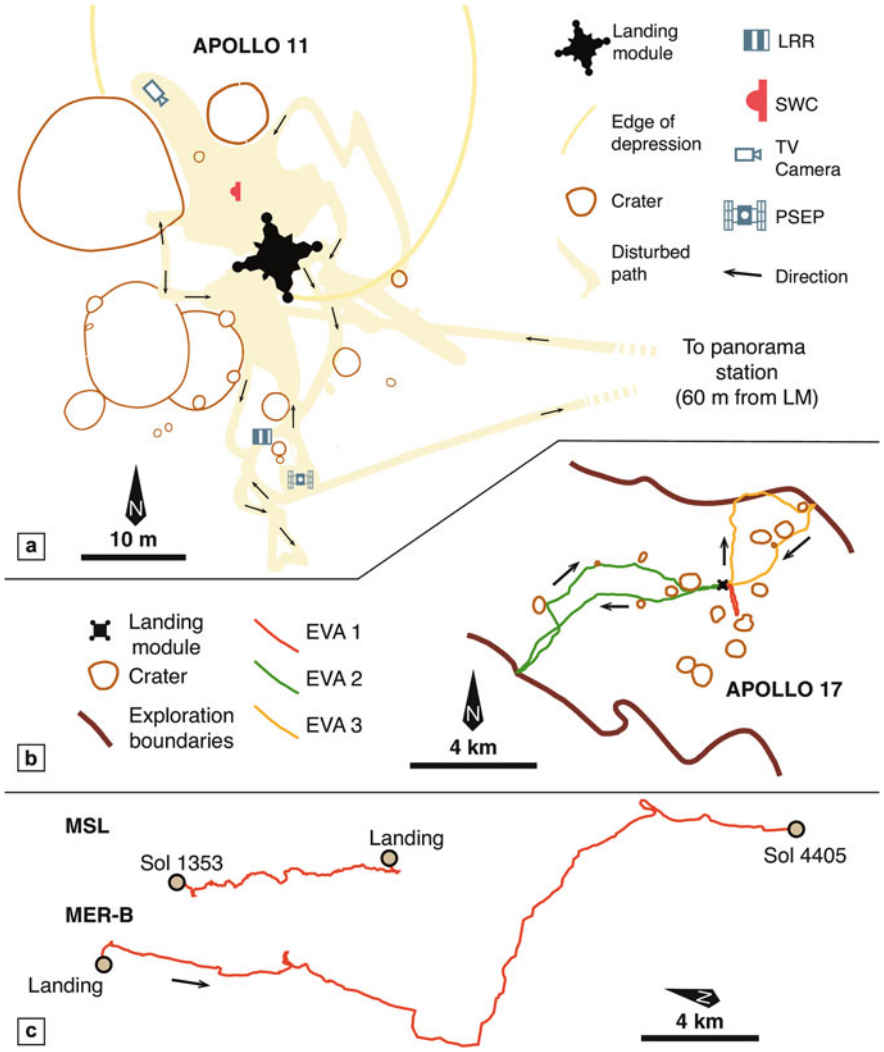

Fig. 5.5 Rover traverses on Moon and Mars, displaying the evolution of Apollo traverse during EVAs from Apollo 11 to 17. (**a**) Simplified traverse map from Apollo 11, with indication of main physiographic features, main direction of movement (on foot) of the astronauts and location of selected experiments, such as LRR, SWC, PSEP. (**b**) Apollo 17 traverse map with a much larger area explored; (**c**) scale of MSL Curiosity first 1353 Sols of exploration in comparison to the distance traveled by MER-B Opportunity after 4405 Sols of primary and several extended missions. Source: (**a**)–(**c**) NASA

potential resources at the surface or in the near subsurface of the Moon (Chap. 15). The first rover that returned to the Moon after Apollo and Luna was China's Chang'e-3 rover Yutu in 2013 (Fig. 5.2).

5.4 Venus

Venus as a target for geological exploration is naturally suited, being so similar in overall size to Earth, relatively close and relatively difficult to observe in detail due to its dense atmosphere and cloud cover, and yet, so visible in the night sky and familiar to most ancient civilisations in Latin. Also, Venus was the first Solar System body beyond the Moon to be reached by a spacecraft fly-by, with NASA Mariner 2 in 1962. Its robotic study turned out to be complex and with its own odds to the planetary exploration *superpowers* (US and USSR): Unlike for Mars, Soviet spacecrafts, entry probes and landers scored better mission success rates than American ones. In addition to orbital missions (Chap. 3) capable of penetrating the thick cloud cover, e.g. with SAR, the most relevant geological platforms on the ground were the Venera landers.

The specific requirements for ground truth on Venus are related mostly to the challenges it creates to optical Remote Sensing, essentially with an atmosphere opaque at visible wavelengths and at most optical ones (Chap. 3), as well as the extremely harsh surface conditions (more than 90 bar pressure, temperatures of around 450°K). Those challenges were both technical regarding the design of spacecrafts able to withstand such conditions (Fig. 5.6), as well as scientific, regarding the characterisation of the surface composition and age with limited remote sensing options. As it is often the case in Planetary Science, what is a limit or handicap in one domain turns out to be an asset in another: the atmosphere opacity and the visibility of Venus' cloud top rather than surface provides observational opportunities for studying its atmospheric dynamics. Environmental constraints to deal with on Venus are essentially its very high surface pressure and temperature, the optical opacity of its atmosphere, the very strong winds characterising its atmosphere (*superrotation*). The barren surface imaged by the Venera landers is dominated by volcanic materials with variable degrees of weathering and alteration (Fig. 5.7). Robotic ground truth on Venus has been achieved with atmospheric probes and static landers. Soviet Venera landers were limited in number (Fig. 5.7) and lifetime, limited by few hours based on battery power and, mostly, environmental harshness, as well as loss of contact with the relay orbiter for communications to Earth. Venera landers evolved through the years in their design, but shared some basic features, such as the quasi-spherical geometry of the pressure vessel (Fig. 5.7a), hosting most experiments and analytical facilities, as well the impact ring at their base (Fig. 5.6), in order to soften the impact on the solid surface. In general, Venera landers were rugged, nevertheless short-lived: their operational lifetime did not exceed a couple of hours. Experiments hosted on Venera (Fig. 5.6b, c), in addition to surface imaging (Fig. 5.7) include investigations on the chemical and isotopic composition with mass spectrometers, gamma-ray spectrometers, as well as gas chromatographs. Measurements of the mechanical properties of the soil were also present, although not always successfully completed.

Future platforms on Venus' surface suitable for geological exploration can be classical static landers or orbiters, as well as balloons. Several concept studies have

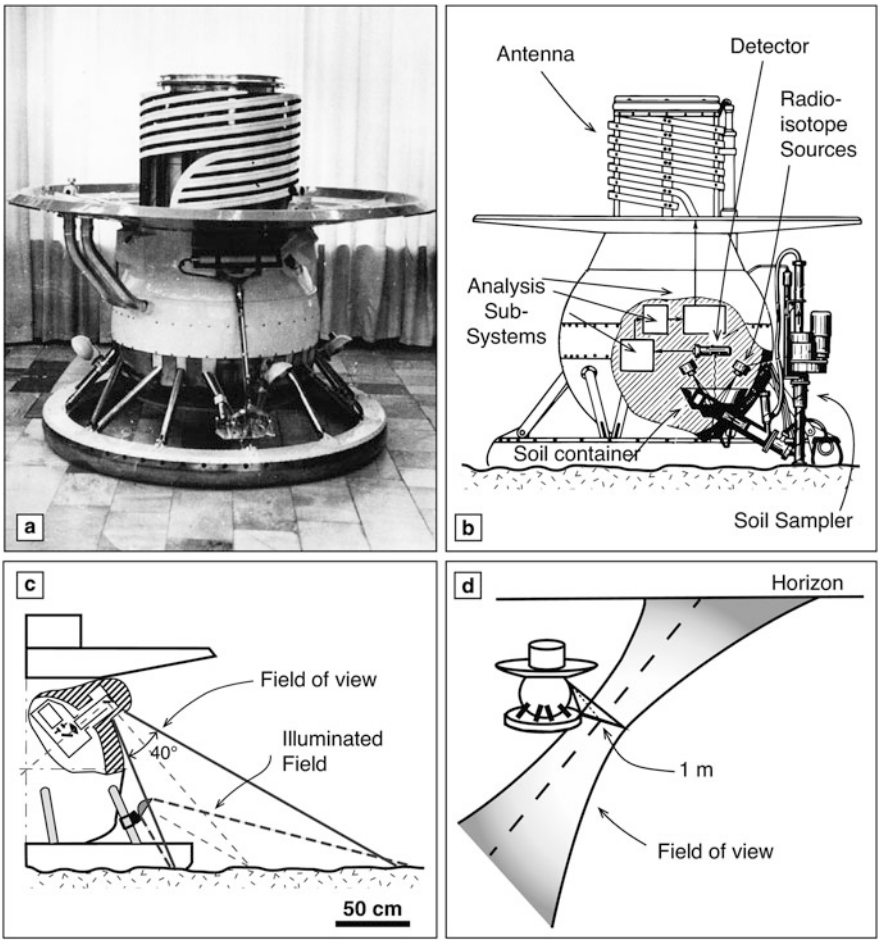

Fig. 5.6 Sample of Soviet VENERA landers with their typical impact ring at their base of round-shaped spacecrafts: (**a**) Venera 9 model. (**b**) Venera 13 simplified scheme and partial cut-out, including the sample acquisition mechanism. Not all subsystems are indicated (**c**) Imaging geometry of Venera TV cameras. (**d**) Ground range of the field of view of Venera TV cameras, as in Fig. 5.7. Source: (**a**) NASA NSSDC. (**b**) Adapted from Surkov et al. (1984). (**c**) Adapted from Florensky et al. (1977)

been carried on balloon explorations and more audacious sailing rover concepts. Regardless the exploration strategies and technologies, the geology of Venus, especially its elusive past, would deserve another close look (Chap. 11), as much of its past is completely unknown. More capable, durable landers with increased mobility and life span would certainly improve our knowledge of Venus' geology. In addition, geophysical measurements would shed light on the interior structure and state of activity of the planet (Chap. 11). Surface nominal lifetime of several tens

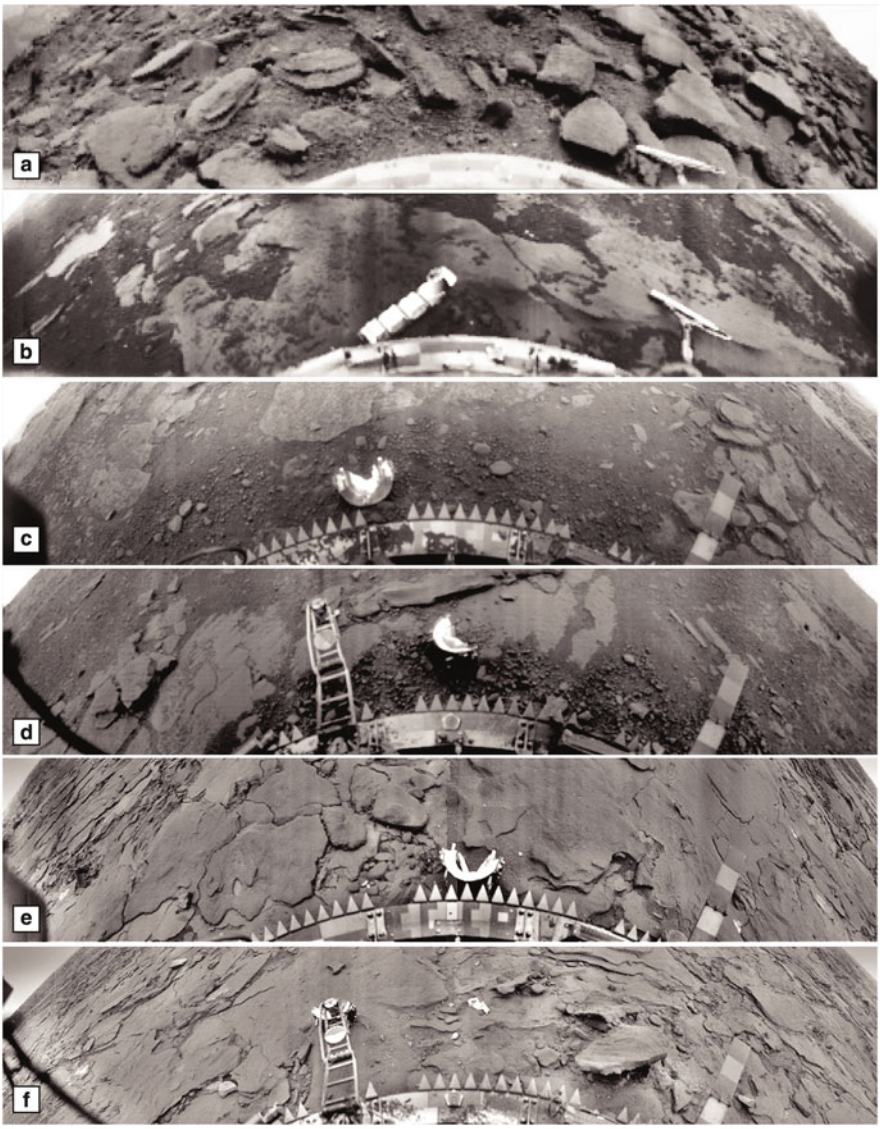

Fig. 5.7 Compilation of surface views on Venus as observed by the USSR Venera landers: (**a**) Venera 9. (**b**) Venera 10. (**c**) Venera 13A, panorama. (**d**) Venera 13B, panorama. (**e**) Venera 14A, panorama. (**f**) Venera 14B, panorama. Source: (**a**)–(**f**) Courtesy of Russian Academy of Science, Sasha Basilevsky

of days and several hundred meters of range would be comparable to the original nominal mission duration of the Mars Exploration Rovers, exceeded by more than one order of magnitude in the case of MER, a case most likely not repeatable on Venus.

5.5 Mars

Due to its geological and environmental relative closeness to Earth, Mars is the most suitable Solar System target for exploration with landers and rovers. Its gravity is in between that of the Earth and the Moon, and the design of some of the landers (e.g. landing gear) is somewhat similar to that of the Moon (Fig. 5.8a–c). The scientific drivers for such exploration are very diverse, ranging from reconstructing the geological evolution of Mars vs. other terrestrial planets, the search for fossil or extant life, as well as the study of resources and risks related of human temporary or permanent settlements (Chap. 15), in perspective. Search for water occurrences and Mars' past and present habitability as a whole (Chap. 14) are both independent drivers as well as topics closely linked to those mentioned above.

In all those cases, ground truth is a key requirement as well as an important validation tool for the extensive remote sensing coverage on Mars. Sample return would crown ground truth activities on Mars as it did for the Moon. The natural next step for obtaining ground truth would be human exploration, but the complexity, cost and risks are much higher, although recent developments in commercial space exploration might help bridging this gap in the near future.

Mars' surface environment is by far less extreme than that of Venus (high pressure and temperature), the Moon (virtually no atmosphere, extreme temperature range) or Mercury (extremely challenging thermal environment), but it does have its own difficulties regarding exploration. Unlike in the case of the Moon, the travel time of information of 12 min on average between Earth and Mars makes real-time commanding unfeasible, thus autonomous or semi-autonomous operations are performed by all lander missions. Landing in the first place is complicated by a dynamic atmosphere, with time-variable characteristics (e.g. dust contents, physical parameters, optical properties and related trajectory and communication problems). Several missions have been lost, even after successful orbit insertion and deployment. Once landed, missions powered by solar panels need to deal—again—with dust deposition from the atmosphere. Fortunately in the mid 2000s NASA's Mars Exploration Rovers (MER) demonstrated that natural dust lifting events linked to the passage of dust devils (Chap. 9) serve as a natural dust removal process, increasing eventual lost efficiency of dusty solar panels. Temperatures on the surface can reach and cross 0°C but they typically vary around several tens centigrades below zero. At medium to high latitudes seasonal frost sublimation and condensation occurs and low areas where surface pressure is higher (e.g. Valles Marineris, Hellas) are often affected by substantial cloud or dust cover. Landing is typically too complicated for highly elevated areas on Mars, such as the large volcanoes located on the Tharsis bulge (Chap. 8) due to the reduced atmospheric braking effect on entry probes, and active braking would require too much fuel.

Thus, in addition to scientific constraints and requirements, engineering constraints for landing mission include among others topographic height, surface roughness at multiple scales, as well as dust cover, or illumination conditions for both thermal and energy availability (e.g. through photovoltaic panels) for the

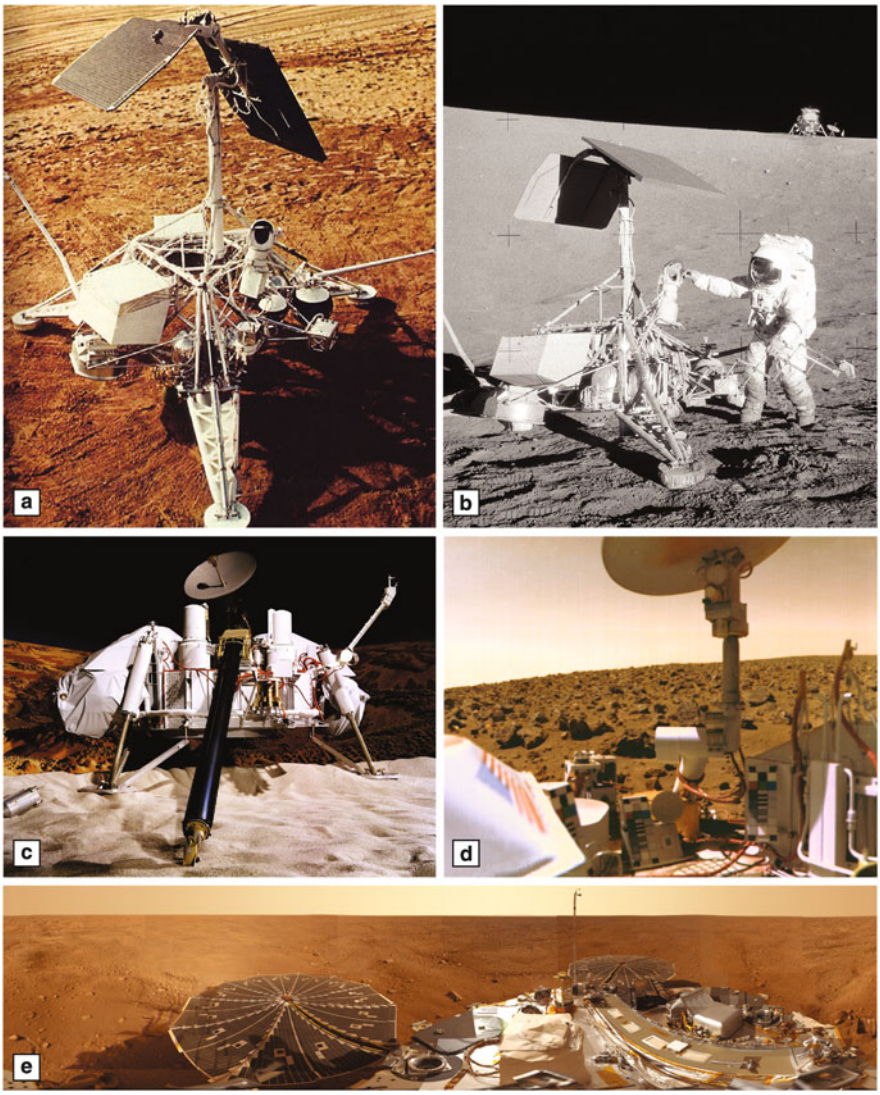

Fig. 5.8 Comparison between Lunar Surveyor lander and Viking and following Mars landers. (**a**) Model of Surveyor-III on a terrestrial beach; (**b**) actual Surveyor-III lander visited by Apollo 12 astronauts; (**c**) model of a Viking lander with sampling arm in the foreground; (**d**) view from the Viking 1 lander-mounted cameras; (**e**) self-portrait of the Phoenix Mars lander from the mast, see also Fig. 5.9. Source: (**a**, **b**) NASA NSSDC. (**c**, **d**) NASA. (**e**) NASA Planetary Photojournal image PIA13804

lander/rover. Depending on the exact scientific mission specifications and spacecraft layout parameters their values vary considerably.

Ground truth on Mars so far used static landers (Viking 1 & 2 landers) as well as rover with an egress platform as a base (Mars 3, Sojourner, MER) and lately a more complex rocket-powered so-called *sky crane* delivering a massive rover (MSL), shared also with Mars 2020. The range of lander types used on Mars, compared with the unmanned spacecraft sent to the Moon a decade earlier is depicted in Fig. 5.5. One can roughly compare the range of distances covered by rovers on Mars and the Moon. Mars 3 probably moved only for an extremely short distance before malfunctioning, Sojourner moved around a hundred meters, MER-A (Spirit) some 8 km, MER-B (Opportunity) around 50 km over more than a decade, and MSL has been travelling tens of km. Moon rovers, both robotic ones (Lunokhod) and human-driven ones (Apollo) covered distances of several tens of kilometres (Fig. 5.5). The scientific value of such range capabilities is substantial, as it allows to reconstruct the space-space or space-time (Chap. 2) local and regional geological variations, as well as the large-scale to global effects or stratigraphic markers (Chap. 11). This my be accomplished partially by high-resolution orbital data (Chap. 3) or by surface-based information (gradient of certain compositional parameters as measured in-situ, or the geometry of geological bodies).

In the decades to come, human exploration, if actually achieved, will most likely start on a temporary, relatively short-term, basis, although far longer than what has been done initially on the Moon due to obvious differences in distance, required travel time, higher complexity of life-support systems, and overall challenges of interplanetary transfer. Human ground truth operations will necessarily rely heavily on robotic aid for transportation, communication and logistical aspects, as well as for scientific and resource-related measurements (Chap. 15). They might get equipped with hardware probably not too different from that of advanced rovers with powerful analytical capabilities (such as MSL Curiosity), with a variable need for autonomy.

Experiments conducted on the surface and strictly related to ground truth of orbital observations are mainly those constraining elemental or mineralogical composition (e.g. remote-sensing spectroscopy on both orbiters and rovers vs. mineralogical or compositional data from in-situ observations, see Chap. 3). In that respect ground-truth experiments on Mars are therefore not different from those that would be carried out on the Moon.

Early rover-based exploration of Mars involved relatively small robots of fractions of a meter in length. On the other hand supporting landers allowing rover egress, were much larger, with footprints of up to few meters.

The very first lander deployed on Mars was the Soviet Mars 3 (location in Fig. 5.3), also featuring a miniature tethered rover which could not be deployed itself due to an early malfunction of the mother lander, which collected and transmitted data only for fractions of minutes before ceasing to work. The Viking landers (Fig. 5.8c–d) were delivered to the surface from Mars orbit by the Viking 1 and 2 orbiters, respectively. The landing-site selection was carried out based on the analysis of data collected by the orbiter themselves, during a period of about 1 month

before landing. Operation of the RTG-powered Viking landers lasted few years which was even longer than the operation time of the Viking orbiters themselves. The Viking landers had a comprehensive payload aimed at characterising the surface and—as a prominent objective—to seek for signs of extant life. The payload included seismometers, imaging cameras, biology experiments and chemical analytical instruments such as gas chromatographs and mass spectrometers, as well as experiments for determining soil properties.

After almost two decades of relatively unsuccessful exploration attempts on and around Mars, the Mars Pathfinder mission succeeded to deploy a small rover on the floor of an ancient large-scale outflow channel (Chap. 9). The small rover Sojourner had a very limited payload, initially just APXS as contact instrument for determining in-situ elemental composition as well as a body-mounted camera. The main camera was hosted on the fixed lander and egress platform, monitoring the progress of the untethered rover in the vicinity of the landing site. The overall operation radius of Sojourner was several meters around the lander, providing communication from the rover to Earth.

Static landers with a very similar design and different fate (Mars Polar Lander and Phoenix) were developed with specific targets on the high latitude of Mars (Fig. 5.8e).

In 2003, NASA's Mars Exploration Rovers were launched and made a substantial technological and scientific step in exploring the geology of two locations on Mars: Gusev Crater and Meridiani Planum (Fig. 5.3). The two sites were rather different. Meridiani Planum is dominated by sedimentary rocks of mainly eolian to evaporitic nature while Gusev Crater consists mostly of volcanic materials with a wide range of water alteration of those. Gusev crater is largely covered by basalts, overlying most likely former paleolake sediments. The lack of enough detail at the time of landing-site selection and geomorphic indicators for a palaeolake in Gusev resulted in a relative surprise when volcanic material was found on the surface. This highlights the importance of landing-site selection and the challenges of the process, due to both data availability at the moment of investigation as well as limited knowledge of extraterrestrial local geology. Endurance, mobility and time are also of importance. Well after the nominal mission, MER rover Spirit reached the foothills of more ancient terrains at Gusev (Columbia Hills) that were not covered by later volcanics.

The payload of MER rovers allowed for morphological as well as chemical and mineralogical characterisation of surfaces and very shallow subsurface materials (Figs. 5.9, 5.10). It included panoramic to microscopic imagers (Fig. 5.9a), contact experiments (Fig. 5.10a), as well as chemical and compositional instruments.

Unlike larger static landers such as Phoenix, MER rovers did not host an internal analytical laboratory for processing and analysing collected samples, nor tools suitable for drilling or excavating beyond rind (Fig. 5.10c). Access to unaltered material is granted thanks to rock surface coating removal devices, both for MER (RAT) and MSL (DRT, Fig. 5.10d). Later rovers (MSL and the Mars 2020 rover) share some of the capabilities but largely exceed them in both number and variety of experiments (Fig. 5.10).

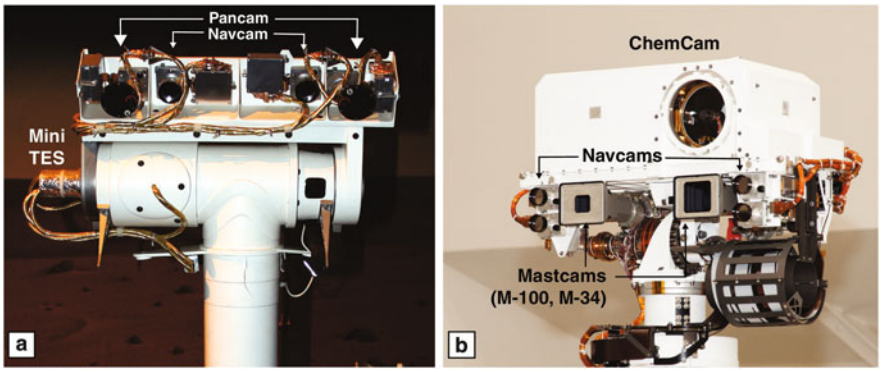

Fig. 5.9 Examples of instruments mounted on robotic masts for Mars surface landers and rovers: (**a**) MER mast with both panoramic and navigation cameras; (**b**) MSL mast-mounted experiments, including the LIBS experiment ChemCam, navigation and panoramic cameras. Source: NASA

The integration of different rover-based remote sensing and in-situ and contact experiments allowed for an appropriate geological characterisation of encountered outcrops and terrains. The multiply extended mission also permitted to cover large distances and comparably broad geologic traverses (Figs. 5.3, 5.5). Such range of observation scales, from microscopic scales to outcrop sizes and towards even larger extends, is ideal to investigate using high-resolution orbital imagery and local rover observations. This will ultimately allow for correlations and transfer of knowledge to other locations (Chap. 2, 3).

In 2011 NASA's Mars Science Laboratory (MSL) Curiosity was launched to Mars to land in Gale Crater. It is the largest rover to date with a scale significantly larger than that of MER (Fig. 5.11). MSL operates irrespectively of daytime illumination, unlike MER, due to its nuclear-powered RTG. This increases productivity in data collection when sunlight is not needed, e.g. while drilling or performing operations without the need of visible imaging of outcrops or panoramic views.

The amount and quality of data used to select and characterise MSL's landing site is superior to any previous lander on Mars. This is also thanks to the favorable string of successful, long-lasting orbital mission providing a plethora of data as well as relay possibilities from global and targeted landing-site observations. Not all future missions to Mars might enjoy the same data coverage and diversity for sites not covered by archived data.

The scientific payload of MSL is both evolutionary and revolutionary compared to that of MER. It includes imaging experiments of different kind, both mast-mounted (Fig. 5.9b) and arm-mounted (Fig. 5.10e,f), as well as a suite of support tools for removing dust cover/rind, drilling, collecting sample and delivering to the internal analytical laboratory with powerful capabilities (Chap. 3), including several experiments carrying out chemical and mineralogical measurements. It also includes

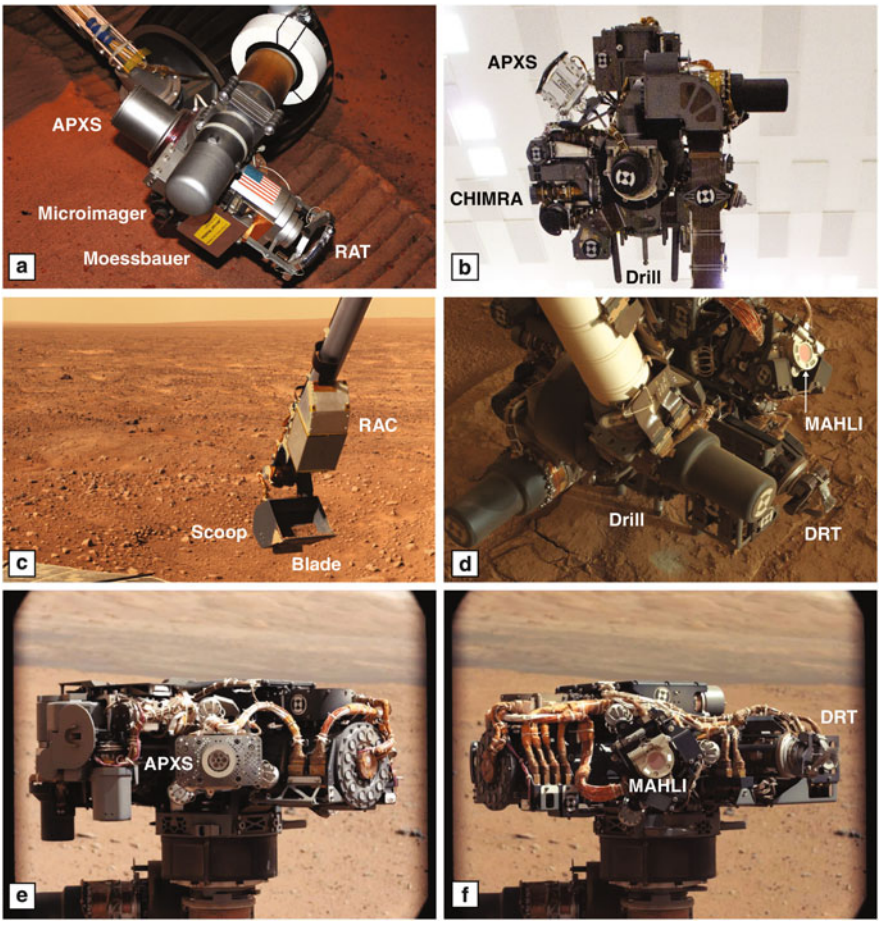

Fig. 5.10 Examples of instruments mounted on robotic arms for Mars surface landers and rovers: (**a**) model of MER robotic arm with contact experiments and surface preparation/abrasion tool; (**b**) view of the partially integrated MSL robotic arm with few experiments and support tools for surface processing and sample acquisition indicated; (**c**) Phoenix lander's robotic arm with scoop and arm-mounted camera, provided by LED, for imaging samples collected and later sent to the lander analytical laboratory; (**d**) MSL robotic arm flight model on Mars with the driller positioned on the surface; (**e**) picture of one side of the robotic arm of MSL with APXS imaged by Mastcam; (**f**) rotated view of MSL robotic arm with dust removal tool and hand lens-like imager (MAHLI) covered by its protection lid

LIBS paired with a camera (ChemCam) capable of both imaging targets analysed through laser ablation as well as long-range imaging, somewhat similar to binocular observations in the field.

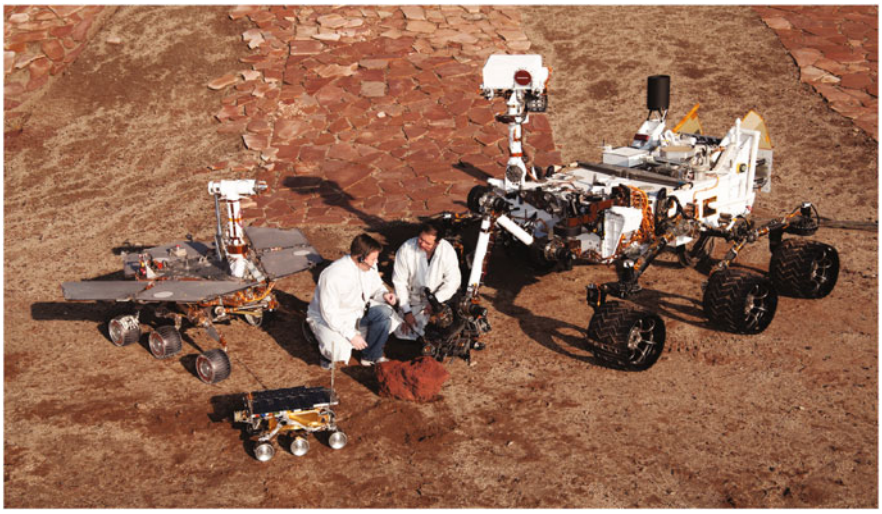

Fig. 5.11 Comparison of the last two decades of Mars rover models to scale: *from left to right* Mars Exploration Rover (2003+), Mars Sojourner (1997), Mars Science Laboratory (2011+). Source: NASA

The geological exploration of Gale and the traverse from its floor towards its central mound Mount Sharp is a multi-kilometer, multi-year endeavour and it makes use of orbital assets in addition to large-scale panoramic information and local in-situ data. Facies analysis and palaeoenvironmental reconstructions are possible, with both geometrical and compositional/geochemical data. Even absolute dating of collected materials has been performed using mass spectrometers on board the rover. A simple example of the very many panoramic views of outcrops collected during the long traverse so far in Gale crater is provided in Fig. 5.12, where images as well as in-situ measurements were collected. Those panoramas are obviously not the sole tool for interpreting local geology and its position in the overall evolutionary scenario of Mars (Chap. 11), but as demonstrated with MER and perfected with MSL, integrated use of independent datasets allows for more reliable geological interpretations, such as the nature of sedimentary rocks and their attribution to a certain sedimentary environment or system.

The capabilities of upcoming rovers, such as Mars 2020, share some overall design with MSL and will allow for sample caching and hosting of a local analytical laboratory. The presence of a complex analytical laboratory is also shared by biosignature-seeking rovers such as ESA ExoMars 2020. Caching of samples on a large, well-equipped rover requires a suitable choice which is strongly dependent on available information at multiple scales, as well as sufficient time for analysing data and selecting, choosing and eventually discarding a sample. This is contrary

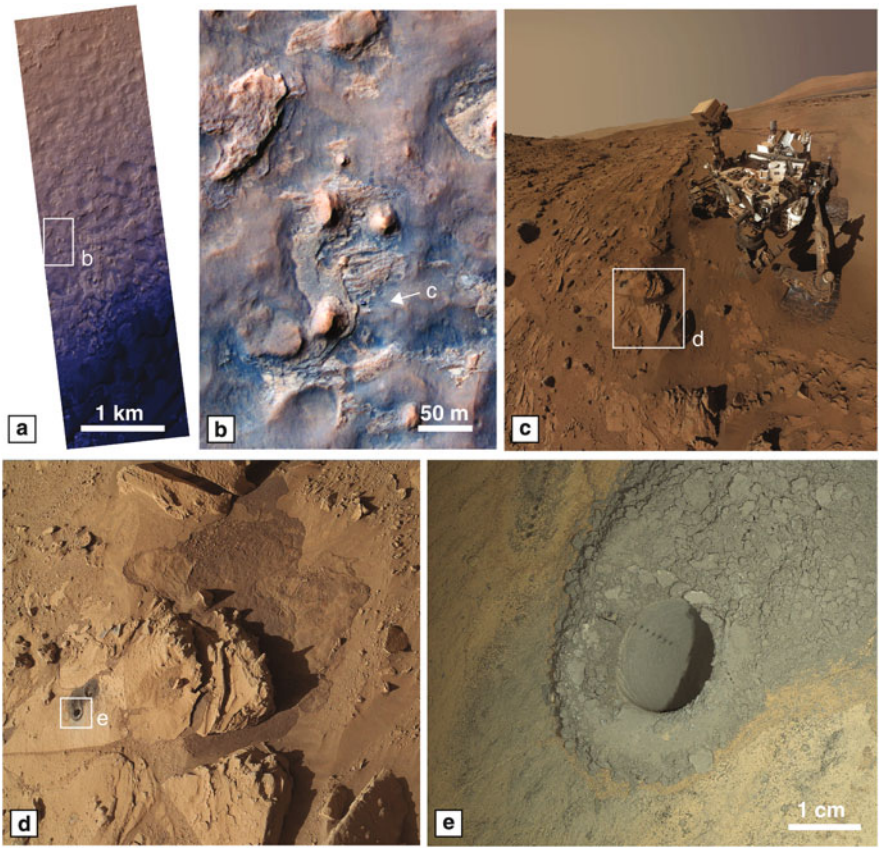

Fig. 5.12 Multiple scales of observations are available at selected landing site locations: nested examples from the Kimberley outgroup and Windjanadrill site on Gale Crater. The Kimberley rocks include several meters thick sedimentary rocks ranging from fine sandstone to conglomerate, interpreted to record an ancient fluvio-deltaic depositional system; (**a**) subset of HiRISE orbital acquired during rover operations; (**b**) close-up of the rover location; (**c**) self-portrait of MSL from MAHLI imagery collected between April 27 and May 12, 2014; (**d**) MAHLI image showing both the Windjana and a small preparatory drill hole, later filled in with cuttings from the main one; (**e**) Windjana drill hole showing aligned markings of ChemCam LIBS analyses along the inner wall of the drill hole itself. Field-based panoramic photos, matched with additional remote sensing and in-situ experiments are the base for geological interpretation. Similar outcrop mapping/line-drawing or context imagery is available from human platforms such as with NASA Apollo or robotic moon landers, until the recent Chang'e 3 rover Yutu. Source: info from Le Deit et al. (2016). (**a**, **b**) NASA/JPL/University of Arizona. (**c**)–(**e**) NASA/Caltech/JPL/MSSS

to the alternative *grab-and-go* approach that is partially blind and poorly contextualised and which is more suited for short-lived fast-paced sample return missions (Chap. 3).

Excursion 5.1 (Choosing the Right Place and Samples)
The choice of suitable landing sites is a challenge in itself given the costs and complexity of landing missions. It involves a more or less democratic selection process, with spacecraft and mission safety being the first priority, since a site with high scientific value but without a surviving spacecraft would lead to zero science return. Site are typically selected after a series of community-driven workshops that help to shortlist a limited number of landing sites based on scientific merit and safety constraints. The final landing site is typically chosen buy the space agency though for private missions the overall process might be slightly different (Fig. 5.13).

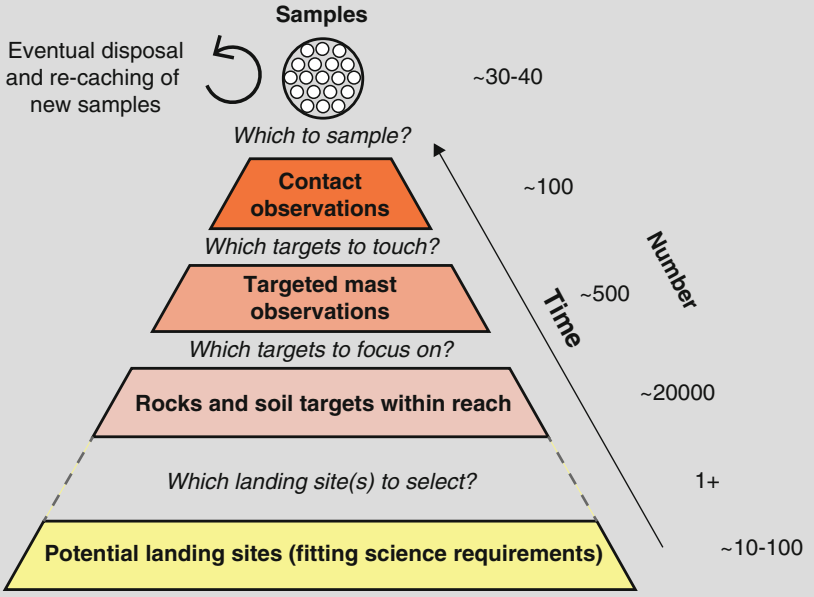

Fig. 5.13 Schematic process of sample selection, given a certain landing sites fitting scientific requirements, based on progressively more detailed observations leading to final selection. Samples, e.g. for sample return, such as MSR, are progressively characterised, collected and possibly discarded in favour of new ones. Source: modified from McLennan et al. (2012)

(continued)

Excursion 5.1 (continued)

Leaving aside *grab-and-go* sample return, very ambitious and complex tasks such as those of Mars Sample Return (MSR) are based on careful and informed selection of landing site(s). Once one or more potential landing sites are selected, the selection of outcrops to be imaged can be guided by existing knowledge (e.g. from orbital data), as well from long-range observations on the ground. Observations accumulate through time and are—based on the judgement of experiment and mission teams—used to *zoom in* to outcrops, beds, individual rocks or smaller-scale subdivisions with increased spatial resolution, contact experiments and extractions of samples for laboratory measurements to be carried out inside the rover. The capability of caching samples is for example embedded in NASA's Mars 2020 rover. The actual number of samples needed to address a given scientific question might vary depending on the type of objective as well as the needed details.

The number of observation leading to the collection of such a number of samples would need to follow a pyramidal approach, based on a very wide range of possible rocks/samples/specimens at reach (tens of thousands), remotely imaging hundreds to thousands, performing contact measurements on tens to hundreds, selecting then the most representative, yet keeping an option to discard some of them later if more compelling samples are to be found.

5.6 Ground Truth of Small and Remote Objects

Relatively small, often irregular solid bodies are scattered across the Solar System. Few of them are satellites, such as Mars' moons Phobos and Deimos, and most of them are either located in the Asteroid Belt or, further away, in the Kuiper Belt and the Oort Cloud, beyond Pluto. Pluto itself is classified as a dwarf planet and visited for the very fist time by a fly-by of NASA New Horizons (Chap. 13). The most accessible ones are the one close to Earth, so-called *Near Earth Asteroids (NEA)*, relevant for both their potential hazard (Chap. 7) and even more for their importance as (near) future source for economically important materials (Chap. 15). Exploring them geologically with landers poses largely different challenges and drives radically different requirements.

5.6.1 Small Bodies: Asteroids and Comets

The exploration of Asteroids and comets, often very irregular in shape, is a challenge for orbiters and even more for landers, having to manouver on a very low gravity

environment and with surfaces that can be extremely heterogeneous at small spatial scale, as demonstrated by missions such as JAXA Hayabusa and ESA Rosetta to asteroids Itokawa and comet 67P/Churyumov-Gerasimenko, respectively.

Regardless of the exact purpose of the lander, the timescale of landing and landing operation has a higher pace than in the case of Moon or Mars and it can be counted in minutes to hours, rather than weeks to years. Hayabusa delivered few miniaturised hopping landers (Minerva) that missed the small target asteroid, and the mother spacecraft itself, as designed, performed sample return during a planned quick touchdown, which turned out to last longer than expected, before lifting off again. The sampling operation on missions to Phobos would last slightly longer, but most of the characterisation of the site in all those cases would have to be done by the orbiter mission and not before, since most small bodies are not imaged with high resolution before an actual spacecraft visits them.

Also the size of landers tends to be smaller, as well as the number and complexity of their payload. In the case of a cornerstone mission such as Rosetta with a rather large and complex lander such as Phylae, the number of experiments can be large, although the time of operation might be limited. This limitations might also be because of the difficulty of operations and unexpected characteristics of the target comet, e.g. with long-range relocation and bouncing (Fig. 5.14a) of the Philae lander itself (Fig. 5.14b).

Bouncing or hopping are sometime not seeked actively (Philae), but in some cases they are a handy way to relocate on a low-gravity body with opportunity for more than one observation point, thus increasing the scientific return. Such approach is shared by micro-landers such as Minerva (Hayabusa, Hayabusa 2) and MASCOT (Fig. 5.14c), where mobility is achieved with moving masses inside the spacecrafts.

The range of geometrical complexity can be exemplified in Fig. 5.14a and Fig. 5.14e: Philae hosted a very wide range of experiments, some similar to those on planetary landers, some different: e.g. APXS, soil properties, mass spectrometry, a magnetometer, an imaging experiment, some of which performed together with the orbiter (e.g. radio sounding, CONSERT). MASCOT, sensibly smaller, hosts imaging camera and spectrometers and a radiometer, as well as a magnetometer (Fig. 5.14e).

5.6.2 Outer Solar System and Water Worlds

Outer Solar System Satellites in the Jupiter and Saturn Systems display an extremely wide range of surface characteristics (Chap. 12). Some are characterised by icy crusts, with or without an atmosphere, some with liquid bodies (Titan), some with subsurface oceans (Europa, Enceladus, Titan, etc.). So far the only ground truth achieved is by the Huygens probe carried from the NASA-ESA Cassini-Huygens mission (Fig. 5.15a), which, back in 2004, returned data from both the atmosphere and surface of Titan. While the spacecraft slowed in Titan's atmosphere, its descent imager (DISR) covered the gap in resolution and scale between fly-by observations by Cassini orbiting in the Saturn system and close-range lander observations.

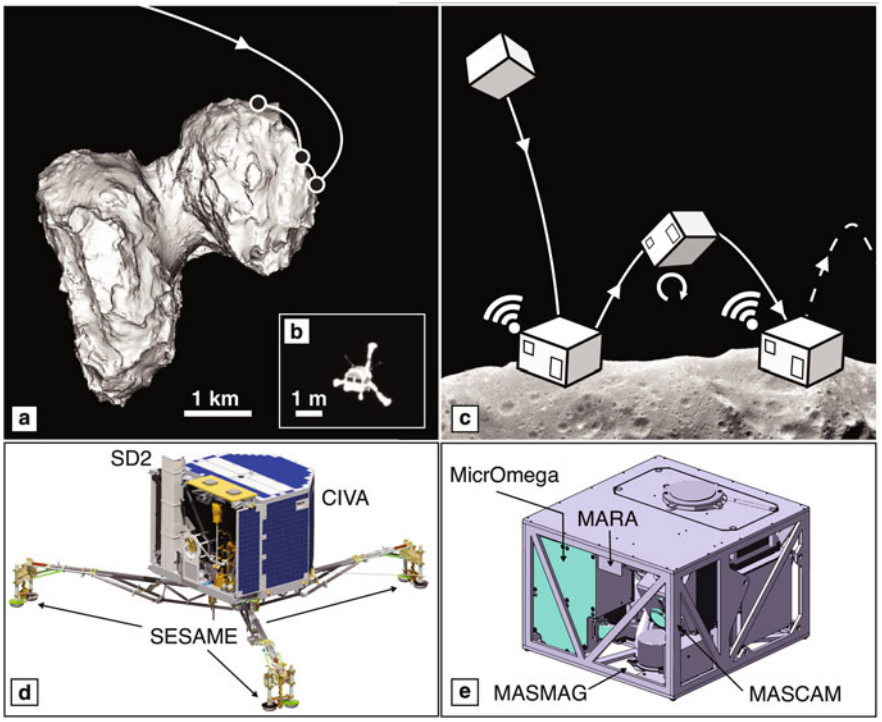

Fig. 5.14 Examples of landers on small bodies, characterised by low gravity: (**a**) cartoon (not to scale) depicting the landing and bouncing process of Rosetta's lander Philae on comet 67P/Churyumov–Gerasimenko on the side view of an artificially illuminated shape model; (**b**) Philae en route to the comet pictured by Rosetta OSIRIS; (**c**) cartoon (not to scale) describing the operations of an asteroid hopping lander such as DLR MASCOT: upon landing, phases of data collection are separated by phases of relocation; (**d**) Philae lander with partial indication of experiments and tools, such as the driller (SD2), the imaging experiment (CIVA) and a set of instruments (SESAME). The entire lander is about 2 m wide; (**e**) MASCOT asteroid lander, on board JAXA Hayabusa-2, capable of re-orienting itself and hopping, in order to collect data at different locations. Source: (**a**) ESA/Matthias Malmer. (**b**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA. (**d**) DLR, courtesy Stephan Ulamec. (**e**) DLR, courtesy Tra Mi Ho

Huygens itself was designed to withstand a landing on both liquid and solid surfaces. Its payload included several atmospheric experiments as well as imaging and surface instruments for addressing acoustic sounding or soil properties. The combination of remote sensing and contact experiments with descent imaging and orbital data provided a consistent view of the only other planetary object beyond Earth with a global, large-scale hydrological cycle and erosional as well as depositional features (Chap. 9) on a non-rocky substrate.

Future missions with increased mobility are desirable. For the Moon and Mars such mobility is achieved with rovers, in the case of Titan, with such a vast amount of hydrocarbon seas, a floating and/or submersible vehicle could achieve

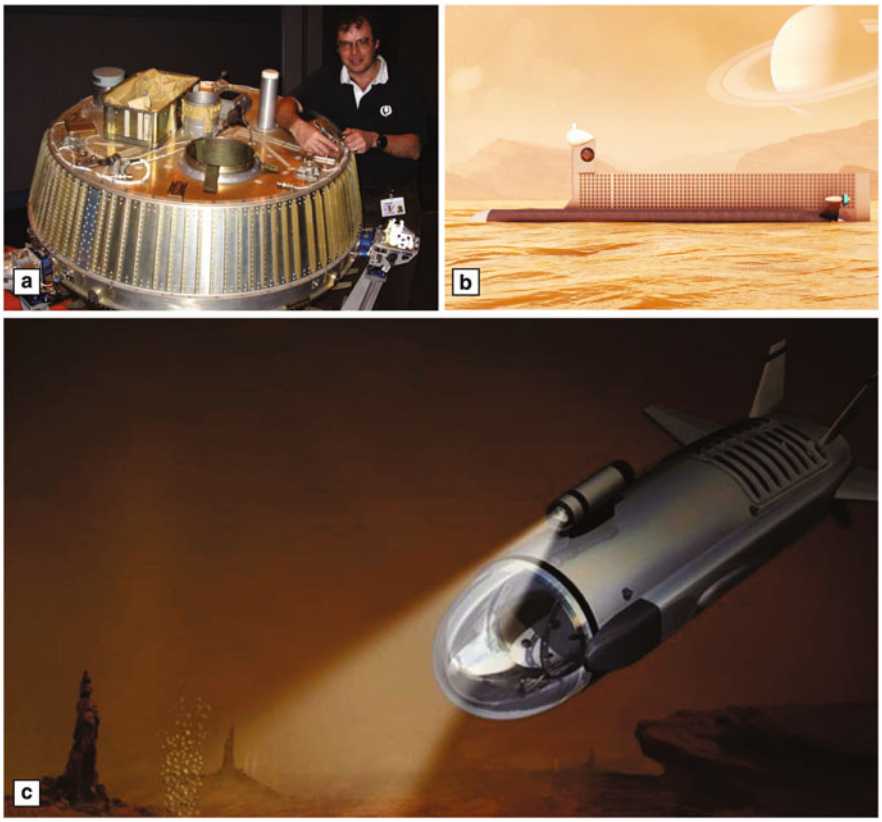

Fig. 5.15 Ground truth platforms for ocean world exploration: (**a**) ESA-NASA Huygens, delivered by Cassini on Titan. The lander, capable of floating, collected data both during atmospheric entry and on the surface; (**b**) concept of a floating-submersible vehicle exploring hydrocarbon seas on Titan; (**c**) artistic view of a submersible in Titan's seas, similar to ocean exploration ROVs, potentially similar to a subsurface ocean explorer in the outer Solar System. Source: (**a**) Courtesy Ralph Lorenz. (**b**)–(**c**) NASA

that (Fig. 5.15b, c). Much larger water bodies are most likely present in several other moons orbiting Jupiter (e.g. Europa) and Saturn (Enceladus). Landing on their surface or accessing their subsurface are very different challenges, both qualitatively and quantitatively. Reaching subsurface oceans covered by a considerably thick icy crusts is still beyond reach, but it would possibly share something with current Earth ocean exploration as well as Titan mare exploration platform designs (Fig. 5.15c). Beyond Saturn, no ground truth exists and only flybys were performed until today. One of them is NASA's New Horizons mission to Pluto (Chap. 13), which is also set to explore more Kuiper Belt objects. Spacecraft designed to land in such remote worlds are still far ahead.

5.7 The Future of Ground Truth

Ground truth operations have evolved enormously in terms of technology in the last several decades, although the range of explored distance of modern robotic platforms still scales similarly to that of Apollo 11–17 (Fig. 5.5). Sample collection and caching capabilities during Apollo and Luna times were quantitatively very large with over 300 kg of material collected over 6 Apollo missions.

Some drivers for exploration include the search for ancient and present life as in the case of Mars or the icy satellites (Chap. 14). Others are related to the evaluation and exploitation of extraterrestrial resources (Chap. 15). In particular the exploitation of planetary resources becomes increasingly important and endeavors develop beyond the fence of space agencies towards commercial actors. The actual success of those endeavours is obviously still to be seen.

The additional ground truth aspect for Mars and the Moon in particular is related to long-term human permanence, which poses challenges beyond resources alone. The geological relevance of such long-term options is the unique Earth-like, geologic mapping and data collection set of task that would naturally come. Crafting a geologic mapping on Earth is an exercise involving a large number of people, a consistent number of collected samples, diverse analytical facilities as well as time, measured between months and years for large projects. Lander and rovers have increasingly powerful capabilities relevant to geology, but the perspective of a long-term, human-robotic exploration on a terrestrial planet such as Mars would have a much larger and scientific return.

As it has been the case for the Moon since early Apollo times there is a trend towards further integration of human and robotic exploration capabilities, in terms of autonomous mapping and geological characterisation of surface and subsurface materials. In recent years Earth's geologic mapping is embedding the use of robotic platforms such as drones, UAV, in planning and executing geological and geospatial mapping tasks. This can be envisaged to develop further on Earth and to be applied, possibly expanded, also to Mars human-robotic geological mapping activities. The stricter constraints on mobility of astronauts related to accessibility and safety a well as life support limitation on Mars would benefit even more than on Earth from reconnaissance and robotic-aided fieldwork. Several analogue operational or geologic mapping campaigns, locally or remotely operated, have been carried out by either space agencies, scientific, industrial or private entities.

Some radically different platforms, either robotic or supporting human exploration are likely to emerge and to be used in actual missions, from terrestrial planet hoppers over large-scale slopes, to tethered robots in very steep areas, up to zoomorphic legged robots, the variety of possibilities to carry out geological observation is extremely large. The same goes for *water worlds*, where on the other hand platforms might resemble more those of marine geologic exploration. In the case of very remote targets miniaturisation of experiments and subsystems will be on the instrument side, while in the case of human exploration larger-

scale geological analysis instruments might be more common. On the long run unexpected technological developments will change this, too.

Ground truth even on relatively close objects such as the Moon is increasingly a collaborative effort with the need of several teams. Permanent settlements on the Moon (*Moon village*) will offer unique occasions for detailed geological mapping sharing the cost and effort among several national or international partners producing a cumulative results in terms of samples, data and overall geological understanding. Those data, archived, curated and made available to the scientific community and the public at large might typically be used for decades after their collection, as Apollo has shown.

The next few decades might see a renewed impulse towards manned exploration of the Moon and Mars, likely preceded by sample return campaigns, to a variable extent. As much as Apollo made the difference in understanding the geologic evolution of the Inner Solar System as a whole, human-robotic exploration and even long-term permanence on Mars would be extremely valuable. Human ground truth on Mars, more than on earlier Moon exploration, would need to deal with planetary protection issues (Chap. 14), whose severity is not yet fully definite.

The additional physical long-term heritage of ground truth in the Solar System consists of the landed hardware itself: robotic geological explorers left on the surface of most visited planetary bodies can in fact last relatively intact for almost geological time scales, although planets with more reactive or corrosive atmospheres as well as with an active volcanism or hydrological cycle (e.g. Venus, Io, Titan) would not preserve as much landed robotic artifacts.

Starting from Apollo, the study of such heritage is the subject of Space Archaeology. Thanks to very high-resolution imaging (MRO HiRISE, LRO LROC), a large part of both successful and failed landers on Mars and the Moon were identified on orbital imagery. So far, a visit of landed hardware by human explorers has been achieved only in the case of Apollo 12 landing in the vicinity of Surveyor III (Fig. 5.8).

Take-Home Messages

The age of planetary exploration started in the late 1950s immediately after the success of the USSR's Sputnik launch into orbit in 1957. The Moon was the first target and the USSR succeeded to hard-land the first spacecraft, to image the farside for the first time and to soft-land a spacecraft on the Moon until the mid 1960s in the course of the Luna program.

The 1960 were marked by a number of very successful mission to the Moon: The Luna program, Ranger, Surveyor, Lunar Orbiter and, finally, Apollo which lasted from 1961 to 1972 culminating in 6 missions to land on the Moon. After 1972, planetary exploration a caesura with only little activity for a decade. Lunar exploration was revived in 1994 with the Clementine and Lunar Prospector

programs. It took until the first decade of the twenty-first century that not only the US, but also Japan, China and India successful sent spacecraft to the Moon.

Venus was visited many times by the Venera program in the early years of planetary exploration. The spacecraft and probe design had to cope with the extraordinary challenging atmospheric conditions on Venus.

Mars has been studied in much detail since the 1970s. The Viking mission, Pathfinder as well as Mars Global Surveyor can be considered milestone missions until the late twentieth century. Europe succeeded to send the Mars Express orbiter to Mars and new high performance rovers (MER, MSL Curiosity) investigated the planet and changing the picture of Mars.

Resource extraction will be a major driver in upcoming decades and apart from the Moon and Mars, small bodies such as asteroids and comets will be a major target. New robotic instrumentation will need to be developed to perform autonomous operations to accomplish investigations.

Further Readings

Ball, A., Garry, J., Lorenz, R., Kerzhanovich, V.: Planetary Landers and Entry Probes, Cambridge University Press, Cambridge (2007)

Bibring, J.-P., Rosenbauer, H., Boehnhardt, H., Ulamec, S., Biele, J., Espinasse, S., Feuerbacher, B., Gaudon, P., Hemmerich, P., Kletzkine, P., Moura, D., Mugnuolo, R., Nietner, G., Pätz, B., Roll, R., Scheuerle, H., Szegö, K., Wittmann, K.: The Rosetta Lander (Philae) investigations. Space Sci. Rev. **128**, 205 (2007)

Campbell, B.A.: Planetary geology with imaging radar: insights from earth-based lunar studies, 2001–2015. Publ. Astron. Soc. Pac. **128**(964), 062001 (2016)

Crawford, I.A., Anand, M., Cockell, C.S., Falcke, H., Green, D.A., Jaumann, R., Wieczorek, M.A.: Back to the moon: the scientific rationale for resuming lunar surface exploration. Planet. Space Sci. **74**(3), 3–14 (2012)

Crumpler, L.S., Arvidson, R.E., Squyres, S.W., McCoy, T., Yingst, A., Ruff, S., Farrand, W., McSween, Y., Powell, M., Ming, D.W., Morris, R.V., Bell, J.F., Grant, J., Greeley, R., Des-Marais, D., Schmidt, M., Cabrol, N.A., Haldemann, A., Lewis, K.W., Wang, A.E., Schröder, C., Blaney, D., Cohen, B., Yen, A., Farmer, J., Gellert, R., Guinness, E.A., Herkenhoff, K.E., Johnson, J.R., Klingelhöfer, G., McEwen, A., Rice, J.W., Rice, M., deSouza, P., Hurowitz, J.: Field reconnaissance geologic mapping of the Columbia Hills, Mars, based on Mars Exploration Rover Spirit and MRO HiRISE observations. J. Geophys. Res. **116**(E00F24) (2011). doi:10.1029/2010JE003749

Garvin, J.B., Head, J.W., Zuber, M.T., Helfenstein, P.: Venus: the nature of the surface from Venera panoramas (1984). J. Geophys. Res. **89**(3381), 3381–3399

Gorman, A.C., Beth Laura O'Leary The archaeology of space exploration. In: Graves-Brown, P., Harrison, R., Piccini, A. (eds.) The Oxford Handbook of the Archaeology of the Contemporary World, pp. 409–424. Oxford University Press, Oxford (2013)

Grant, J.A., Golombek, M.P. Grotzinger, J.P., Wilson, S.A., Watkins, M.M., Vasavada, A.R., Griffes, J.L., Parker, T.J.: The science process for selecting the landing site for the 2011 Mars Science Laboratory. Planet. Space Sci. **59**(1114), 1114–1127 (2011)

Grotzinger, J.P., Crisp, J., Vasavada, A.R., Anderson, R.C., Baker, C.J., Barry, R., Blake, D.F., Conrad, P., Edgett, K.S., Ferdowski, B., Gellert, R., Gilbert, J.B., Golombek, M., Gömez-Elvira, J., Hassler, D.M., Jandura, L., Litvak, M., Mahaffy, P., Maki, J., Meyer, M., Malin,

M.C., Mitrofanov, I., Simmonds, J.J., Vaniman, D., Welch, R.V., Wiens, R.C.: Mars Science Laboratory mission and science investigation. Space Sci. Rev. **170**(5), 5–56 (2012)

Ho, T.-M., Baturkin, V., Grimm, C., Grundmann, J.T., Hobbie, C., Ksenik, E., Lange, C., Sasaki, K., Schlotterer, M., Talapina, M., Termtanasombat, N., Wejmo, E., Witte, L., Wrasmann, M., Wübbels, G., Rößler, J., Ziach, C., Findlay, R., Biele, J., Krause, C., Ulamec, S., Lange, M., Mierheim, O., Lichtenheldt, R., Maier, M., Reill, J., Sedlmayr, H.-J., Bousquet, P., Bellion, A., Bompis, O., Cenac-Morthe, C., Deleuze, M., Fredon, S., Jurado, E., Canalias, E., Jaumann, R., Bibring, J.-P., Glassmeier, K.H., Hercik, D., Grott, M., Celotti, L., Cordero, F., Hendrikse, J., Okada, T.: MASCOT – The mobile asteroid surface scout onboard the Hayabusa 2 Mission. Space Sci. Rev. **208**(1–4), 339–374 (2017). doi:10.1007/s11214-016-0251-6

Hurtado, J.M., Young, K., Bleacher, J.E., Garry, W.B., Rice, J.W.: Field geologic observation and sample collection strategies for planetary surface exploration: insights from the 2010 Desert RATS geologist crew members. Acta Astronaut. **90**, 344 (2013)

McLennan, S.M., Sephton, M.A., Allen, C., Allwood, A.C., Barbieri, R., Beaty, D.W., Boston, P., Carr, M., Grady, M., Grant, J., Heber, V.S., Herd, C.D.K., Hofmann, B., King, P., Wilson, M.G., Mangold, N., Ori, G.G., Rossi, A.P., Raulin, F., Ruff, S.W., Sherwood Lollar, B., Sy, S.: Planning for Mars returned sample science: final report of the MSR End-to-End International Science Analysis Group (E2E-iSAG). Astrobiology **12**(3), 175–230 (2012)

Soderblom, L.A., Tomasko, M.G., Archinal, B.A., Becker, T.L., Bushroe, M.W., Cook, D.A., Doose, L.R., Galuszka, D.M., Hare, T.M., Howington-Kraus, E., Karkoschka, E., Kirk, R.L., Lunine, J.I., McFarlane, E.A., Redding, B.L., Rizk, B., Rosiek, M.R., See, C., Smith, P.H.: Topography and geomorphology of the Huygens landing site on Titan. Planet. Space Sci. **55**, 2015 (2007)

Squyres, S.: Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet. Hachette Books, New York (2005)

Surkov, Y.A., Barsukov, V.L., Moskalyeva, L.P., Kharyukova, V.P., Kemurdzhian, A.L.: New data on the composition, structure, and properties of Venus rock obtained by Venera 13 and Venera 14. J. Geophys. Res. **89**, B393–B402 (1984). doi:10.1029/JB089iS02p0B393

Ulamec, S., Biele, J., Bousquet, P.-W., Gaudon, P., Geurts, K., Ho, T.-M., Krause, C., Lange, C., Willnecker, R., Witte, L.: Landing on small bodies: from the Rosetta Lander to MASCOT and beyond. Acta Astronaut. **93**, 460 (2014)

Part II
Processes and Sources

Chapter 6
Meteorites

Ansgar Greshake and Joerg Fritz

6.1 Introduction

Planet Earth is continuously accreting extraterrestrial material. Every day, millions of tiny particles totaling to about 100 tons are collected as Earth orbits around the Sun. Most of the material is dust with grain-sizes ranging from less than a micrometer to only few millimeters. Larger objects fall less frequently and their relative abundances are expressed in size-frequency distribution diagrams. With increasing size these extraterrestrial objects arriving on Earth are termed interplanetary dust particles (IDPs), micrometeorites, meteorites and impact crater forming projectiles. This terminology is specific for Earth as it depends on the density of the Earth's atmosphere that controls the deceleration and fictional heating of objects arriving with various velocities, entry angles, initial size and types of material. The encounter velocities vary between 11.2 and 72 km/s and are constrained by the maximum velocity above which an object would leave the Solar System, and the minimum velocity due to gravitational acceleration by Earth.

The smallest of these extraterrestrial particles are indirectly visible in the sky and they are known before sunrise and after sun dawn as the zodiacal light and at midnight as *Gegenschein*—sun light scattered by <1 to few μm-sized dust grains concentrated in the ecliptic plane near the Sun and extending to the asteroid belt. These dust particles are debris of asteroid collisions or liberated from comets during their passage through the inner Solar System. Due to their high surface-mass ratio allowing them to radiate heat very quickly, IDPs can be trapped in Earth's

A. Greshake (✉)
Museum für Naturkunde, Berlin, Germany
e-mail: ansgar.greshake@mfn-berlin.de

J. Fritz
Saalbau Weltraum Projekt, Heppenheim, Germany
e-mail: joerg.fritz@kino-heppenheim.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_6

atmosphere without being melted or completely vaporized. Many IDPs collected in the stratosphere are fluffy aggregates of loosely consolidated mineral grains with a highly porous texture.

Most of the larger, mm-sized material entering Earth's atmosphere melts or vaporises by frictional heating within 1–2 s in ~60 to 130 km altitude. Such meteors appearing as bright flashes at the sky can either be sporadic or occur as showers. Meteor showers are strongly periodic and result from Earth's yearly passage through the stream of small particles left from certain comets or rarely from asteroids. Meteor showers are often named after the star constellation in which they appear to originate (radiant). In famous showers such as the Perseids (in August, from comet 109P/Swift-Tuttle), the Leonids (in November, from comet 55P/Tempel-Tuttle), and Geminids (in December, from asteroid 3200 Phaethon) the number of meteors per hour increases from typically 10–60 to more than 1000.

Depending on their initial size, entry angle and velocity a small amount of mm-sized particles reaches altitudes of 10–30 km without being vaporised and falls as micrometeorites on Earth's surface. Some of them remain completely unmelted retaining primordial characteristics while others are partly or completely melted. Micrometeorites have been collected from ocean and terrestrial sediments but most successfully and least altered from polar deposits including Antarctic blue ice and Greenland snow. Most of them are similar to carbonaceous chondrites in mineralogy and chemical composition and likely represent asteroidal material.

6.2 Meteorite Falls and Finds

While the vast majority of extraterrestrial material reaches Earth as dust and micrometeorites, falls of larger meteorites are relatively rare. Find statistics and direct observations made by automatic cameras may indicate a rate of some 500 meteoroids entering Earth's atmosphere each year of which about ten are finally recovered as meteorites.

Once a meteoroid reaches the outer fringes of the atmosphere it will be decelerated by collision with air molecules in the atmosphere until it finally loses its cosmic velocity at an altitude of ~100 km. Meteoroids weighing more than ~10,000 kg will not be effectively stopped by the atmosphere and impact onto Earth's surface with at least part of their cosmic velocity. During deceleration the surface of the meteorite is rapidly heated to temperatures above the melting point of most silicates between 1773 and 2273 K leading to significant mass loss (roughly 30–60%) due to ablation. At about 2273 K the ablated material and in particular the atmospheric gas along the path of the meteoroid become ionized resulting in a brilliant fireball. During its hypersonic flight the meteoroid compresses the air in front and thus generates shock waves that can result in heavy detonations, similar to those produced by supersonic aircrafts.

Ablation and fragmentation during the meteoroid's flight through the atmosphere determine the final shape of the meteorite. Stony meteoroids often break up

into many smaller fragments due to drag forces. If fragmentation occurs high in the atmosphere, most edges will be rounded and the surface smoothened during ablation. Often small pits termed *regmaglypts* can be observed at the surface indicating some turbulence during melting. Iron meteoroids have a much higher strength and are, thus, more rarely fragmented than stony objects. They usually have highly irregular shapes with deep smoothened depressions resulting from severe melting. Irons that exploded upon impact on Earth often look shrapnel-like with sharp edges. Both, iron and stone meteorites will develop a thin fusion crust during quick cooling. Below the few mm thick fusion crust the meteorite remained cold due to low thermal conductivity of the rock and even delicate structures such as hydrated phyllosilicates and magnetic properties with low curie temperatures are preserved.

Collecting meteorites one generally distinguishes between witnessed meteorite falls and meteorite finds. Currently there are about 64,700 meteorites listed in the data base of the *Meteoritical Society*, of which only 1,304 are falls, all others are finds (September 2017[1]). Once officially classified, meteorites are given unique names usually referring to a nearby landmark.

Whereas finding a meteorite in densely populated regions or areas with lots of vegetation is rather unlikely, especially Antarctica and the Sahara turned out to be ideal places for meteorite search. Since 1969 more than 40,000 meteorites have been recovered from various Antarctic blue ice fields by scientific expedition. Meteorites found in Antarctica are usually only very little weathered and can have high terrestrial ages of up to 900,000 years. Each find is carefully documented and classified. Names assigned to these meteorites contain the find location, the year of the expedition and a specimen number. For example, the meteorite *ALH 84001* is the first specimen classified from those recovered by the 1984 expedition to the Allan Hills ice field.

More recently, a large number of meteorites were found in the Northern Sahara and in the Arabic desert of Oman. Those with no exact documentation of the find circumstances are named after so-called dense collection areas followed by a number (e.g., Northwest Africa 952). Meteorites found in the deserts are often severely weathered giving them a rusty appearance. Most are also sandblasted, i.e., the original black fusion crust was partly or completely removed.

6.3 Origin of Meteorites

It is widely accepted that the vast majority of meteorites originates from the main asteroid belt a region between Mars and Jupiter, 2–4 AU away from the Sun. The about 2 AU wide region is populated by more than 200 asteroids larger than 100 kilometers, more than 750,000 asteroids larger than 1 km in diameter and probably millions of smaller objects (see Chap. 13).

[1] Source http://www.lpi.usra.edu/meteor/.

Most of our knowledge about asteroids comes from telescope observations, close encounters by space missions, spectral reflectance measurements and most recently investigations of samples returned directly by space crafts. Using reflectance spectra of minerals present at their surface, asteroids have been classified into 3 main groups and about 26 sub-groups of different composition. The main groups are C-type (carbon-rich), S-type (stony), and M-type (metallic). The classification of asteroids also showed that a relation exist between the asteroid's composition and position in the belt, i.e., the inner belt contains those asteroids composed of refractory minerals and only few volatiles while C-type asteroids rich in water-bearing minerals, volatiles and organics dominate the outer belt. Spectral reflectance measurements of asteroidal and meteorite surfaces that yield close matches also allowed to link meteorites to certain types of asteroids. Meteorites are fragments from likely more than 135 different parent bodies that either preserved pristine material from the early solar nebula or document different degrees of aqueous alteration, thermal metamorphism and finally magmatic differentiation.

6.4 Classification of Meteorites

Meteorites are coarsely divided into stony (∼97%), iron (∼2%) and stony iron (1%) meteorites. More precisely and related to their formation process they are classified into *undifferentiated chondrites*, *primitive achondrites* and *differentiated achondrites*. The subsequent detailed classification into different classes and groups basically rests on the chemistry, mineralogy, petrology, and oxygen isotopic compositions of the different types of meteorites (Fig. 6.1).

6.4.1 Chondrites

Chondrites are the oldest and most primitive rocks of the Solar System. They are *cosmic sediments* that predominantly consist of chondrules, refractory objects (Calcium-aluminum-rich inclusions (CAIs) and amoeboid olivine aggregates), FeNi metal and fine-grained matrix. Refractory inclusions, chondrules and FeNi metal formed very early in the solar nebula by high-temperature processes including condensation and evaporation; fine-grained matrix consists of silicates, oxides, sulfides, and carbon that formed at much lower temperatures. Both high and low-temperature fractions of the solar nebula finally accreted to chondritic rocks.

Chondrites are divided into three major classes, *carbonaceous chondrites (C)*, *ordinary chondrites (O)* and *enstatite chondrites (E)*. Further subdivision yields 13 groups of chondrites (Fig. 6.1).

- *Carbonaceous chondrites* (Fig. 6.2) are divided into eight groups: CI, CM, CR, CH, CB, CV, CO and CK. While *C* stands for *carbonaceous*, the second letter

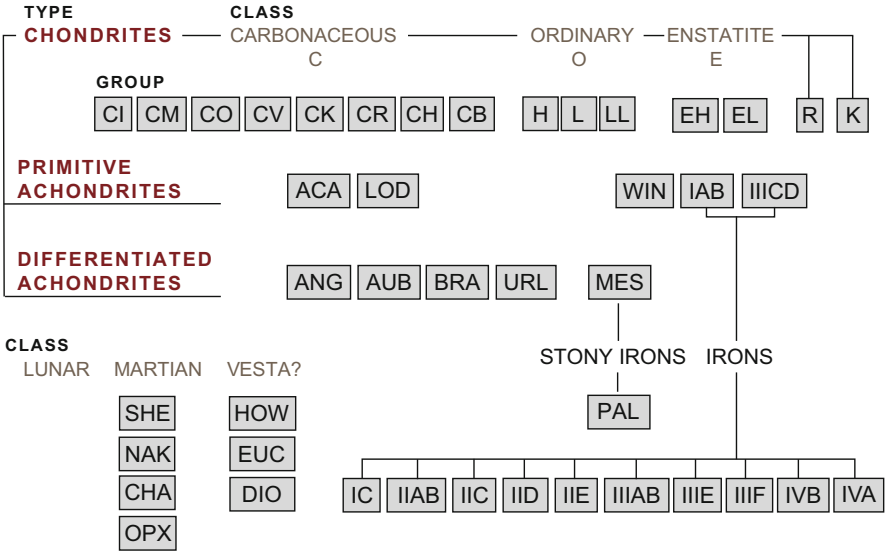

Fig. 6.1 Classification of meteorites. ACA acapulcoites, *LOD* lodranite, *WIN* winonaites, *ANG* angrites, *AUB* aubrites, *MES* mesosiderites, *BRA* brachinites, *URE* ureilites, *SHE* shergottites, *NAK* nakhlites, *CHA* chassignites, *OPX* orthopyroxenite, *HOW* howardites, *EUC* eucrites, *DIO* diogenites, *PAL* pallasites

Fig. 6.2 Sawn surface of the carbonaceous chondrite Allende (CV3). The meteorite is composed of rounded chondrules, irregular shaped whitish Ca−Al-rich inclusions and FeNi metal grains all cemented by a dark carbon-rich matrix. Specimen is 7 cm wide

refers to the typical chondrite fall in the respective group: CI (Ivuna-like), CM (Mighei-like), CR (Renazzo-like), CH (ALH85085-like), CB (Bencubbin-like), CV (Vigarano-like), and CK (Karoonda-like). In the CH group, the *H* refers to high metal abundance and high iron concentration.

- *Ordinary chondrites* are divided into three groups: H chondrites with high total iron contents, L chondrites with low total iron contents, and LL chondrites which have low metallic iron relative to total iron, as well as low total iron contents. Ordinary chondrites are called ordinary because they constitute 87% of all classified meteorites, with H (39%), L (35%) and LL (13%). They, however, appear to represent only 3–4 different parent bodies.
- The class of *enstatite chondrites* contains two groups with different contents of metallic iron: EH have high-iron and EL low-iron contents.
- *R (Rumuruti-like) chondrites* show some similarities with ordinary chondrites but also considerable differences and constitute a separate group.
- *K (Kakangari-like) chondrites* are also distinct from other chondrites and considered a grouplet, i.e. less than five meteorite of this type are known.
- Meteorites that cannot be assigned to one of the above groups are commonly named *ungrouped*.

The different chondrite classes and groups are distinguished from each other by their unique properties including petrographic characteristics (e.g., abundances and sizes of individual components), bulk composition, oxygen isotopic composition, and oxidation state all attesting to the heterogeneity of their source regions within the solar nebula.

6.4.1.1 Chondrite Properties

The *petrographic characteristics* vary considerable between the different chondrite groups. CI chondrites contain virtually no chondrules while in ordinary and enstatite chondrites they can make up to 80 vol% of the rock. Matrix abundances follow exactly the opposite trend. While the largest refractory inclusions are found in CV chondrites, they are most frequently encountered in CO chondrites. Chondrule mean diameters are well-defined for each group and vary from only 20 μm in CH chondrites to 0.9 mm in LL chondrites indicating that each parent body preferentially collected chondrules of a characteristic size range.

The *chemical composition* of a specific type of carbonaceous chondrites, the CI chondrites was found to be very similar to the chemical make-up of the Sun as determined by the atomic adsorption lines (*Fraunhofer Linien*) in the sunlight except for the highly volatile elements H, C, N, and the noble gases He, Ne, Ar, Kr, Xe. The compositional similarities between CI meteorites and the Sun include chemical elements with different properties such as atomic number and mass, first ionization potential, condensation temperature and geochemical behavior suggesting that no physical or chemical element fractionation processes had changed their abundances.

Because CI chondrites appear to represent the unfractionated composition of the Solar System, they are commonly used as reference material.

Compared to CI chondrites the other chondrite groups are less volatile-rich and most nonvolatile elements vary within a factor of 3 relative to CI abundances. The different chondrite groups also show different degrees of depletions in moderately volatile elements. Each chondrite group has a characteristic elemental abundance pattern that allows distinguishing them from the other groups.

Most chondrite groups also show distinct *bulk oxygen isotopic compositions* and plot in unique positions on a three-isotope oxygen plot ($\delta^{17}O$ vs. $\delta^{18}O$). Enstatite and CI chondrites plot on or very near the terrestrial fractionation line (TFL) while all other carbonaceous chondrites fall below it. Anhydrous components of several C chondrites fall along a slope called carbonaceous chondrite anhydrous mineral (CCAM) mixing line. Finally, ordinary and R chondrites form distinct clusters above the TFL. It should, however, be noted that there is considerable overlap of oxygen isotopic compositions for some groups, e.g. EH–EL, type 3 H–L–LL ordinary chondrites, and CV–CO–CK.

Primary solar nebula and secondary parent body processes had significant influence also on the *oxidation state* of each chondrite group as reflected by the ratio of metallic to oxidized iron. Highly reduced chondrites have high abundances of iron in metal and/or sulfides and low abundances of iron in silicates or oxides; highly oxidized chondrites vice versa. Generally, an increasing degree of oxidation is observed in the order enstatite–ordinary–carbonaceous chondrite classes. R chondrites are similar to CO chondrites; Kakangari chondrites are intermediate between E and O chondrites. Among the ordinary chondrites, the oxidation state increases from H to L to LL, as demonstrated by an increase in the Fe/(Fe+Mg) ratios of their olivine and pyroxene. For carbonaceous chondrites the degree of oxidation state increases in the order CB–CH–CR–CO–CV–CK–CM–CI.

To account for the degree of secondary aqueous and thermal alteration that effected chondrites on their parent bodies van Schmus and Wood introduced a classification scheme that assigns certain petrologic types ranging from 1–6 to chondrites. In this scheme, type 3 chondrites represent the most primitive, unequilibrated material. Decreasing numbers indicate aqueous alteration that results in the production of hydrous phases from anhydrous silicates. Type 1 chondrites are more heavily aqueously altered than type 2 chondrites. Numbers from 4 to 6 attest to an increasing degree of thermal metamorphism that took place at temperatures below the formation of partial melts between 400 and 950 °C. During thermal metamorphism the compositions of olivine and pyroxene homogenize and the outlines of chondrules become progressively indistinct as solid state recrystallisation occurred. Type 3 ordinary, CO, and CV chondrites are divided into 3.0–3.9 subgroups, of which 3.0 is the least metamorphosed.

Further classification parameters are the degree of shock metamorphism and the degree of terrestrial weathering. For shock metamorphism (see Chap. 7) the classification scheme relies largely on shock effects observed in olivine and plagioclase. The scale of terrestrial weathering is based on the progressive oxidation of metals, sulfides, and finally silicates.

6.4.1.2 Chondrite Components

Chondrites are mixtures of diverse components that experienced distinct formation histories before accreting to chondritic rocks.

Chondrules

Chondrules (from Greek $\chi o \nu \delta \rho o \varsigma$, grain, seed or cartilage) are mostly sub-mm sized spherical objects with igneous texture that are dominantly composed of the Mg–Fe-silicates olivine and pyroxene and interstitial feldspathic glass (Fig. 6.3). They are unique to chondritic meteorites and unlike any terrestrial rock. According to their igneous texture and spherical shape formation by crystallisation of molten droplets was inferred since their discovery. Chondrules are considered to have formed by melting of dust aggregates in solar nebula at temperatures well above 1773 K. Possible heat sources required for such high-temperature event include propagating shock waves during accretion of material, lightning, bipolar outflows, jets, and magnetic resonances.

Melting of dust aggregates and subsequent solidification of the melt droplets at different temperatures and cooling rates lead to the development of characteristic elemental abundance pattern and textures of chondrules. Compositionally, most

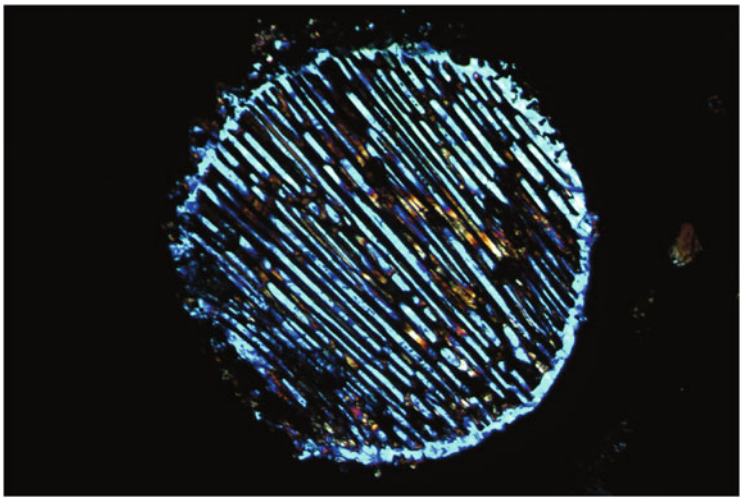

Fig. 6.3 Chondrule from the carbonaceous chondrite Dhofar 1994 (CM2). The barred olivine chondrule consists of sets of parallel running elongated olivine crystals. Transmitted light, crossed polarizers; about 0.5 mm in diameter

chondrules are very similar to CI chondrites with respect to refractory and moderately volatile elements and depleted in volatile and siderophile elements. These differences are either due to reduction and volatile element loss during melting or reflect primary characteristics of the precursor dust material.

Texturally, porphyritic chondrules are distinguished from granular, barred, spherulitic and cryptocrystalline chondrules. Additionally, type I chondrules refer to FeO-poor, reduced chondrules and type II chondrules to FeO-rich, oxidised chondrules.

Many chondrite groups contain compound chondrules composed of two or more conjoined chondrules, which suggests that collisions between molten or partly molten droplets frequently occurred in the chondrule forming regions. Chondrules often show fine-grained rims of mineral dust that accreted onto the chondrules while they were still hot.

Refractory Inclusions (CAIs, Olivine Amoeboids, and *Fremdlinge*)

Calcium-aluminum-rich inclusions (CAIs) are few μm to even cm-sized objects that show diverse outlines from highly irregular to almost perfect spherical (Fig. 6.2). They are especially abundant in carbonaceous chondrites but are also found rarely in ordinary, enstatite and Rumuruti chondrites. CAIs are dominantly composed of oxides and silicates rich in Ca and Al, and to minor degrees also in Mg and Ti. These phases have extremely high vaporization temperatures of >1300 K at ambient space pressure of 10^{-4} bar and are predicted to be the first solids to condense from a gas of solar composition.

Typical mineral phases found in CAIs are (in the order of decreasing equilibrium condensation temperatures): corundum (Al_2O_3), hibonite ($CaAl_{12}O_{19}$), grossite ($CaAl_4O_7$), perovskite, melilite, spinel, diopside, fassaite (Al–Ti-rich pyroxene), and anorthite.

Calcium-aluminum-rich inclusions in CV chondrites occasionally contain metal-rich, opaque assemblages known as *Fremdlinge* (German for *stranger*). Fremdlinge are dominantly composed of FeNi metal, sulfides and Pt-metal grains and are highly enriched in Pt-group elements, such as Re, Os, Ir, Pt, Rh, and Ru. These opaque assemblages are either high-temperature condensates that were later incorporated into their host inclusions or initially formed as homogeneous metallic grains during melting of CAIs and were later subject to exsolution, sulfidation, and oxidation at lower temperatures.

Amoeboid olivine aggregates (AOA) are irregularly shaped, mm to cm-sized lumpy objects dominated by fine-grained Mg-rich olivine grains; some AOAs also contain CAI-like nodules. Amoeboid olivine aggregates are explained as clumps of lower-temperature nebula condensates.

FeNi Metal

FeNi metal grains are a common constituent in most chondritic meteorites and consist of the two phases kamacite and taenite. The low-Ni alloy kamacite can have up to 7.5 wt% Ni in solid solution, while Ni-rich taenite contains 27–65 wt% nickel. In chondrites the two alloys typically occur as a fine-grained intergrowths named plessite.

Matrix Material

Matrix that cements the coarser components in chondrites is highly diverse in mineralogy and chemistry among the different chondrite groups. It consists of intimate mixtures of fine-grained silicates, oxides, sulfides, sulfates, carbonates, FeNi metal and carbonaceous material. Matrix is often affected by hydrous alteration that occurred on the chondrite's parent bodies and converted anhydrous silicates (mainly olivine and pyroxene) into hydrous phyllosilicates (e.g. serpentine, saponite, tochilinite).

Pre-solar Grains

Many primitive chondrites and interplanetary dust particles preserved particles that are older than our Solar System. These pre-solar grains are small, typically few-μm sized individual mineral grains and characterized by isotopic signatures significantly different to the material indigenous to the Solar System.

The list of interstellar phases identified so far includes diamond, silicon carbide, graphite, silicon nitride, several oxides that sometimes occur as sub-μm sized inclusions within other interstellar grains, and also silicates. High precision isotopic measurements of e.g., C, N, O, Si, Ti and Ca, are used to study stellar nucleosynthesis and to determine the stellar sources of the stardust. The available data indicate that many of the interstellar grains formed in asymptotic giant branch stars of 1.5–4 solar masses as well as in certain types of supernovae.

Organic Material

Chondritic meteorites, IDPs and micrometeorites can also contain various types of organic compounds as diverse as insoluble kerogen-like macromolecules and simpler soluble organics ranging from polar amino acids and alcohol containing multiple hydroxyls to nonpolar hydrocarbons. It is likely that the different organics formed by diverse multistep processes involving pre-solar syntheses, solar nebular reactions and parent-body processing. In addition, there is some evidence that isotopically anomalous organic meteoritic material enriched in deuterium and/or ^{15}N originated from interstellar matter or from the outer regions of the Solar System.

6.4.2 Non-chondritic Meteorites

Primitive chondrites testify to the very early processes occurring in the Solar System that are condensation of the first solids from the solar nebula, melting and evaporation, and accretion to form chondritic parent bodies. Very soon after their formation this pristine material was modified to very different degrees by aqueous, thermal and shock metamorphism taking place on the asteroidal bodies.

Progressive thermal metamorphism finally leading to the differentiation of planetary bodies from their primitive chondritic precursors through intense heating is well documented in the suite of *differentiated meteorites*.

Intense heating leading to melting of an asteroidal body of chondritic composition to high degrees requires temperatures sometimes in excess of 1773 K and, thus, enormous and long lasting heat sources. Most likely, differentiated asteroids were heated mainly due to decay of short-lived radionuclides such as ^{26}Al and ^{60}Fe as well as by electrical conductivity heating caused by T Tauri solar wind from the early Sun. Locally, large and frequent impacts caused additional heating and shock melting.

While highly metamorphosed type 6 chondrites have almost completely lost their chondritic texture, chemically they still retain chondritic bulk composition. It is not before temperatures exceed 1223 K that the first metals and sulfides will begin to melt and to separate for the silicate material. The initial stage of this process can be studied in the group of *primitive achondrites* comprised of acapulcoite, lodranites, winonaites and some related silicate inclusions in iron meteorites.

6.4.2.1 Primitive Achondrites

Acapulcoites and Lodranites

Acapulcoites are fine-grained rocks dominantly composed of orthopyroxene, olivine, plagioclase, metal and troilite of approximately chondritic modal abundances. They display a recrystallized texture of 150 to 230 μm equigranular mineral grains often joined at 120° grain boundaries. Rarely relict chondrules are present.

Lodranites show a very similar mineral inventory and also thoroughly recrystallized textures but much larger grain sizes of 540–700 μm compared to acapulcoites. Lodranites are often depleted in plagioclase and troilite relative to chondrites.

Most acapulcoites contain abundant, typically μm to mm-sized FeNi metal and sulfide veinlets which cross-cut all silicate phases suggesting that the silicates were and remained solid during formation and migration of the metal-sulfide-melt. This rather low degree of partial melting attests to peak temperatures just above 1223 K. The bulk rock composition of acapulcoites remained essentially unchanged from their chondritic precursors indicating that the melt was not removed from the rock.

In contrast, lodranites are likely residues of higher degrees of partial melting resulting in formation and loss of a plagioclase-rich partial melt. Peak temperatures for those lodranites depleted in both troilite and plagioclase must have reached about 1523 K.

Winonaites and Silicate Inclusions in IAB and IIICD Iron Meteorites

Winonaites are fine- to medium-grained rocks showing recrystallized achondritic textures of equigranular olivine, orthopyroxene, plagioclase and FeNi metal. They are chondritic in bulk composition and some Winonaites contain relict chondrules and FeNi-metal-sulfide veins.

Silicate inclusions occurring in IAB and possibly also those in IIICD iron meteorites have remarkable similar mineralogy and almost identical oxygen-isotopic composition compared to winonaites suggesting that they originate from the same parent body.

Melting and Differentiation

Textures and compositions of primitive achondrites document early stages of partial melting including formation of FeNi-metal-sulfide-melts slightly above 1223 K and of plagioclase-rich silica melts at about 1523 K.

Asteroids that continue to heat up to temperatures in excess of about 1773 K will be subject of extensive partial melting and efficient segregation of metal to form large cores. Models suggest that core formation requires partial melting of more than 50% for smaller and up to 90% for larger (10 km-sized) asteroids. While the outer asteroid shell consists of an early-formed basaltic crust, it is likely that the mantles of such planetary bodies did not melt completely. Larger bodies, however, may have been capable to reincorporate the crust into the mantle due to melting at the mantle-crust boundary and impact mixing producing a homogeneous magma ocean. Smaller asteroids likely preserved their basaltic crust and maybe even smaller remains of the chondritic precursor material.

Solidification of the magma ocean (Chap. 11) involves equilibrium crystallization followed by segregation of early-formed minerals and subsequent fractional crystallization of the residual melt. A typical large and differentiated asteroid would than consist of an innermost metallic core, followed by a thick layer of olivine that is overlain by harzburgite, orthopyroxenite, or some combination of the two, and finally by an outermost basaltic to gabbroic crust. Larger planetary bodies such as the Moon or the planet Mars can retain their internal heat for a much longer time and produce rocks with a chemically evolved composition. For these planets the Martian and lunar meteorites represent lithological diverse samples of the crust.

Meteorites that attest to these processes and allow insights into the buildup of large asteroids, the Moon and Mars are called *differentiated achondrites*.

6.4.2.2 Differentiated Achondrites

Lunar Meteorites

So far more than 306 different lunar meteorites represent a variety of lithological types including feldspathic and basaltic rocks, granulites, KREEP (potassium [K], Rare Earth Element and Phosphor) rich rocks akin to the Procellarum KREEP Terrane (PKT), monomict breccias, polymict anorthosite-basalt breccias, impact melt rock breccias and regolith breccias. These breccias contain clasts of mafic non-basaltic lithologies such as anorthositic norites, norites, gabbronorites, and troctolites. The compositional diversity and relative abundance of lunar meteorites resemble the relative abundance of related lithologies on the lunar surface as identified by orbital observations. It is therefore assumed that lunar meteorites are a more representative sample suite for the overall compositional diversity and abundance of lunar crustal rocks compared to the Apollo and Luna mission samples.

Martian Meteorites

The suite of Martian meteorites is constituted by about 198 volcanic and plutonic rocks whose chemical and O-isotopic compositions differ from those of other meteorites and suggest their origin from the same differentiated parent body. Young crystallization ages of 180–1300 Myr point to derivation from a planet-sized body and Martian atmosphere found trapped in impact-produced glass inclusions strongly advocate their origin from Mars. Formerly known as SNC (shergottites-nakhlites-chassignite) meteorites, Martian meteorites now comprise basaltic, lherzolitic, and olivine-phyric Shergottites named after Shergotty (fall India, 1865), clinopyroxenites named Nakhlites after Nakhla (fall Egypt, 1911), dunites named Chassignites after Chassigny (fall France, 1815), the orthopyroxenite ALH 84001 and the unusual basaltic breccia Northwest Africa 7034. The Martian meteorite ALH 84001 is the only one representing old Martian crust with an age of about 4000–4500 Myr.

Howardites, Eucrites, Diogenites

Howardites, eucrites and diogenites are commonly termed HED meteorites and constitute a suite of basaltic and plutonic rocks. Identical oxygen-isotopic composition, similar Fe/Mn ratios in their pyroxenes and the occurrence of polymict breccias composed of eucritic and diogenitic material lead to the assumption that HED meteorites originated on the same parent body, presumably 4 Vesta. They, thus, represent a cross-section from mantle to crust of the second-most-massive object in the asteroid belt.

Diogenites are orthopyroxenites dominantly composed of mm to even cm-sized blocky magnesium-rich orthopyroxene grains and minor olivine, calcic plagioclase and chromite. The plutonic rocks are mostly intensely brecciated and display a

characteristic cumulate texture of interlocked mineral grains. Texture and chemical composition indicate that diogenites formed from a fractionally crystallizing magma at greater depth.

Eucrites are pyroxene-plagioclase-bearing basalts that are further subdivided into cumulate eucrites formed by fractional crystallization from a mafic melt, basaltic eucrites representing volcanic rocks that have formed originally as quickly cooled lava flows at the parent body's surface, and polymict eucrites consisting of eucritic clasts, diogenitic fragments, melt rock and glass particles all consolidated in a fine-grained matrix. Determining for polymict eucrites are less than 10 vol% of diogenitic component.

Howardites are very similar to polymict eucrites and distinguished from them by having more than 10 vol% of diogenitic component. Both, polymict eucrites and howardites formed by brecciation and impact mixing at the surface of their parent body.

Meteorites that formed at even greater depths, at the mantle-core boundary of differentiated asteroids are the Pallasites (Fig. 6.4).

Pallasites

Pallasites are stony-iron meteorites composed of Mg-rich olivine set into a matrix of metallic Fe-Ni plus troilite. Olivine and metal are mostly homogeneously distributed but regions dominated either by metal or by olivine are also present. Pallasites are likely derived from the boundaries between mantle and core of differentiated

Fig. 6.4 Slab of the Esquel pallasite consisting of olivine crystals in a matrix of FeNi metal. Specimen is 17 cm wide

asteroids. While the metallic Fe-Ni represents the core material, olivine is the expected lowermost mantle phase. Some compositional, isotopic and textural differences among the pallasites testify to their formation on multiple parent bodies. The cores of large differentiated asteroids are represented by the iron meteorites.

Iron Meteorites

Iron meteorites are largely composed of metallic iron-nickel with minor amounts of troilite, graphite, schreibersite and phosphates. Some of them also contain silicate inclusions.

The metallic portion has typically 5–20% Ni and consists of the two Fe-Ni alloys kamacite (α-FeNi) and taenite (γ-FeNi). Etching a polished surface with nitric acid reveals in many irons a characteristic texture called the Widmannstätten pattern (Fig. 6.5), an oriented intergrowth of kamacite with Ni<6%, and high-nickel regions composed of taenite and several other phases. Kamacite grows precisely oriented along octahedral planes and iron meteorites, which show this structure, are thus called octahedrites. Iron meteorites with more than 15% nickel consist almost entirely of taenite. They show no Widmannstätten pattern and are termed ataxites. Iron meteorites with less than 6% nickel are composed only of kamacite and are named hexahedrites. They show also no Widmannstätten pattern but sometimes develop very fine striations called Neumann lines which are mechanical twins induced by shock metamorphism.

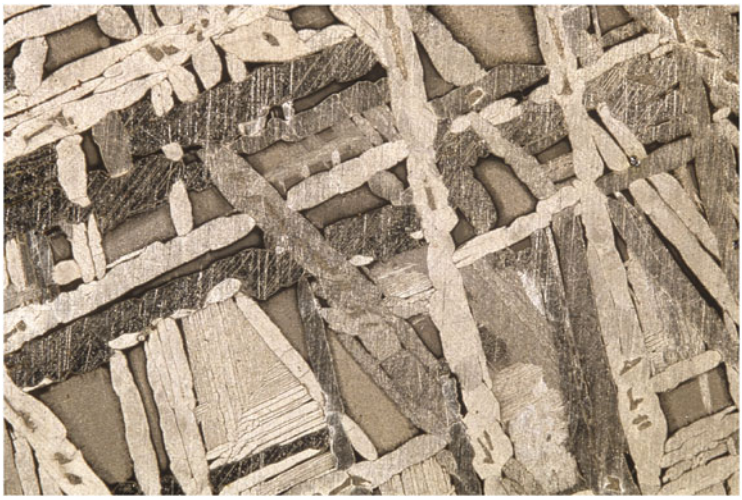

Fig. 6.5 Polished and etched slab of the Carbo iron meteorite (IID) showing pronounced Widmannstätten pattern of kamacite bands and interstitial regions composed of taenite and several other phases. The close-up is 2.5 cm wide

A more genetical classification addressing the formation mechanism of iron meteorites uses the trace element content of the Fe-Ni-metal phases.

Plotting the concentrations of gallium, germanium and iridium versus nickel, 13 different groups accounting for about 86% of all iron meteorites are distinguished. Each of them is designated by a Roman number followed by a capital letter: IAB, IC, IIAB, IIC, IID, IIE, IIF, IIIAB, IIICD, IIIE, IIIF, IVA, and IVB.

The siderophile trace elements Ga, Ge and Ir are used to discriminate between irons because their concentrations reflect the crystallization process of the metallic melt.

During crystallization these elements will strongly fractionate according to their different affinities to liquid and solid Fe-Ni. Their abundance patterns are thus characteristic for distinct metallic melts likely corresponding to different parent bodies.

All iron meteorites that show trace element fractionation formed by fractional crystallization in a metallic core of a differentiated asteroid and are called magmatic iron meteorites.

In contrast, silicate-bearing IAB, IICD and IIE irons have broad ranges in Ni and show only little evidence for fractionation. They are called non-magmatic iron meteorites and formed either in smaller impact produced melt pools or by mixing of metallic and chondritic material, in any case leading to only insufficient metal-silicate separation.

6.4.2.3 Other Achondrites

There are several types of achondritic meteorites whose origin is less clear than those of Martian, lunar and HED meteorites, pallasites and irons reflecting the diverse processes of partial to complete melting of different precursor materials.

Ureilites are ultramafic rocks dominantly composed of olivine and pyroxene. Pyroxene is mostly pigeonite but also diverse assemblages of pigeonite, augite, and orthopyroxenes occur. Ureilites contain significant amounts of interstitial carbon that is either present as graphite or microdiamonds. Uniquely among the achondrites, ureilites have preserved a primitive oxygen-isotopic composition and scatter along the carbonaceous chondrite anhydrous mineral mixing line.

Ureilites are certainly achondrites, but their formation is discussed controversially. While textures, the lack of plagioclase as well as chemical bulk compositions argue for a high degree of igneous processing, preservation of a primitive O-isotope signature seems to contradict long-lasting high-temperature events.

Brachinites are dunitic wherlites dominantly composed of olivine (74–98%) with minor augite, plagioclase, chromite and Fe-sulfides. They show igneous textures but have almost chondritic abundances of lithophile and siderophile elements and of noble gases pointing to only very low degrees of partial melting. Brachinites may have crystallized from a melt of unique composition.

Aubrites, also referred to as enstatite achondrites, are orthopyroxenites that formed under highly reduced conditions. They consist of up to 95% almost FeO-

free blocky enstatite and contain abundant clasts of clearly igneous origin. Aubrites are strongly related to E chondrites but likely originated on different parent bodies.

Angrites are medium- to coarse-grained basaltic rocks and consist mainly of Ca–Al–Ti-rich pyroxene (fassaite), calcium-rich olivine, and plagioclase of essentially pure anorthitic composition. They show porphyritic, ophitic, cumulate, and polygonal-granular textures and frequently contain porous regions and round vesicles. Angrites are highly depleted in sodium and potassium, rich in ferrous oxide and critically undersaturated with silica. Melting experiments suggest that angrites formed by partial melting of CM and CV chondrites under redox conditions where iron metal is unstable.

Mesosiderites belong to the group of stony-iron meteorites. They are polymict breccias composed of about equal proportions of silicates and FeNi-metal plus troilite. The silicate fraction consists of igneous lithic clasts, as well as of orthopyroxene, olivine, and plagioclase mineral fragments. Fe-Ni metal is present as small grains uniformly distributed within a silicate matrix, larger clumps or aggregates, and metal veins. Troilite is abundant in mesosiderites and can be as much as 14 vol%. Mesosiderites formed either by collision of two differentiated asteroids allowing liquid core material to mix with solidified crustal silicates or alternatively by impact-induced break-up and reassembly of a Vesta-like asteroid with a molten core.

6.5 Chronology of the Solar System as Told by Meteorites

The initial nuclide inventory of our Solar System included short-lived radionuclides (parent) that decay into stable isotopes (daughter) with half-live times of less than 10 Myr such as ^{107}Pd/^{107}Ag (9.38 Myr), ^{60}Fe/^{60}Ni (2.16 Myr), ^{53}Mn/^{53}Cr (5.34 Myr), ^{41}Ca/^{41}K (0.149 Myr), ^{36}Cl/^{36}Ar, ^{36}S (0.301 Myr), ^{26}Al/^{26}Mg (1.05 Myr). The former presence of these short-lived radionuclides as testified by isotopic anomalies of the daughter isotopes in CAIs, chondrules and iron meteorites requires a stellar formation very close in time and space to where and when the Solar System formed. This is because the time scale during which planetary systems formed from collapsing molecular clouds is relatively long (10–50 Myr) compared to the lifetimes of these short-lived radionuclides.

The oldest objects of the Solar System are the calcium-aluminum-rich inclusions with an absolute U-Pb age of 4567 Myr. This absolute age constitutes the age of our Solar System and serves as an anchor point to date the relative difference in time between formation of CAIs and the further evolution of the Solar System material. Important dates include the formation of magmatic iron meteorites (<1 Myr after CAI formation), the formation of chondrules (1–3 Myr), thermal metamorphism of the parent bodies of ordinary chondrites (2–4 Myr) and primitive achondrites (1–4 Myr), the magmatic differentiation of asteroids (3–5 Myr), core formation of Mars (2–10 Myr), and the age of the Earth–Moon system (equal core formation ages of 38–128 Myr).

These chronological constraints show that the formation of the planetary bodies and the terrestrial planets in our Solar System occurred in the short time interval of 10–100 Myr starting 4567 Myr ago. The interstellar medium obviously clumped together to an evolved planetary system with a fully developed Mars in less than 50 million years and a fully developed Earth–Moon system about 30–100 Myr later. A time span that is short compared to the geological history of Earth (Chaps. 7 and 11).

Take-Home Messages

Meteorites are extraterrestrial messengers derived from asteroids, the Moon and Mars reporting on the history of the Solar System.

So far about 64,700 meteorites representing more than 130 different parent bodies are classified according to their mineralogical, chemical and isotopic properties.

Meteorites document the very early processes occurring in the Solar System including condensation of the first solids, accretion, aqueous and thermal metamorphism and finally magmatic differentiation of large parent bodies.

The most primitive meteorites allow determining the age and chemical bulk composition of the Solar System.

Differentiated meteorites are used to constrain the time span from the condensation of the first solids to the final formation of the fully evolved planetary system.

Suggested Readings

Ceplecha, Z., Borovička, J., Graham, W., E., Revelle, D.O., Hawkes, R.L., Porubčan, V., Šimek, M.: Meteor phenomena and bodies. Space Sci. Rev. **84**(3–4): 327–471 (1998). doi:10.1023/A:1005069928850

Davis, A. (ed.): Meteorites, Comets, and Planets. Treatise on Geochemistry, vol. 1, p. 756. Elsevier-Pergamon, Oxford (2006)

Fritz, J., Bitsch, B., Kührt, E., Morbidelli, A., Tornow, C., Wünnemann, K., Fernandes, V., Grenfell, J., Rauer, H., Wagner, R. Werner, S.: Earth-like habitats in planetary systems. Planet. Space Sci. **98**, 254–267 (2014). doi:10.1016/j.pss.2014.03.003

Huss, G., Meyer, B., Srinivasan, G., Goswami, J. Sahijpal, S.: Stellar sources of the short-lived radionuclides in the early solar system. Geochim. Cosmochim. Acta **73**(17), 4922–4945 (2009). doi:10.1016/j.gca.2009.01.039

Kleine, T., Touboul, M., Bourdon, B., Nimmo, F., Mezger, K., Palme, H., Jacobsen, S., Yin, Q.-Z. Halliday, A. (2009). Hf-W chronology of the accretion and early evolution of astroids and terrestrial planets. Geochim. Cosmochim. Acta **73**(17), 5150–5188. doi:10.1016/j.gca.2008.11.047

Lauretta, D.S., McSween, H.Y. (eds.): Meteorites and the Early Solar System II. Space Science Series, p. 942. The University of Arizona Press, Tucson (2006)

Lodders, K.: Solar system abundances of the elements. Astrophysics and Space Science Proceedings, pp. 379–417. Springer, Berlin (2010)

Lodders, K., Fegley, B.: The Planetary Scientist Companion. p. 372. Oxford University Press, Oxford (1998)
McSween, H.Y.: Meteorites and Their Parent Bodies, 2nd edn., p. 310. Cambridge University Press, Cambridge (1999)
Norton, O.R.: The Cambridge Encyclopedia of Meteorites, p. 354. Cambridge University Press, Oxford (2002)
Papike, J.J. (ed.): Planetary Materials. Reviews in Mineralogy, vol. 36. p. 864. Mineralogical Society of America, Chantilly (1998)

Chapter 7
Impact Cratering

Thomas Kenkmann and Gerwin Wulf

7.1 Introduction

Impact craters are omnipresent in the solar system: The terrestrial planets and their moons, icy satellites, dwarf planets, and minor bodies of the asteroid belt; all of them are punctured by impact craters of various sizes. There is only one body with a solid surface that does not have them—the Jovian satellite Io. Here the resurfacing rate by volcanic processes is so high that impact craters, once they are formed, rapidly undergo destruction or coverage. Of the terrestrial planetary bodies Mercury and Moon show the most heavily cratered surfaces. Some plains, such as the highlands of the Moon, have even achieved a saturation with craters of a given size. That means for every new crater formed, statistically one old crater of the same size is obliterated. The number and size spectrum of impact craters on a planetary surface relates to the period of its exposure to cosmic bombardment. The heavily cratered surfaces of the Moon, for instance, have ages of 4 Ga and older. The lunar crater production function that relates size and frequency of craters with exposure ages, is the best investigated among all planetary bodies, thanks to a large remote-sensing data base and the absolute dating of the age of some surfaces through the analysis of the 380 kg samples collected during the Apollo and Luna missions. The calibrated Lunar production function could be successfully transferred to other planets such as Mars. For this a number of factors have to be considered including the population of impactors in the orbital range of the planet, their orbital velocity spectrum with respect to the planet, the planet's gravity and strength properties, and the effects of an atmosphere. With the aid of crater densities it is thus possible to determine ages of planetary surfaces. In other words, the geological history of planets and satellites can be unraveled and the tool for this are impact craters.

T. Kenkmann (✉) • G. Wulf
University of Freiburg, Freiburg, Germany
e-mail: thomas.kenkmann@geologie.uni-freiburg.de; gerwin.wulf@geologie.uni-freiburg.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_7

On Mars where exogenic and endogenic processes were more active during its history than on the Moon, resurfacing reduced the total number of exposed impact craters. However, Mars still has an impressive number of approximately 300,000 impact craters larger than 1 km in diameter. On Venus, the crater density is a magnitude less than on Mars. Small craters are entirely missing here due to the shielding effect by a 90 bar dense atmosphere. The reduced number of large craters indicates a planet-wide resurfacing on Venus that occurred on a relatively short timescale. On Earth, the number of discovered craters is currently only 190.[1] A recent study showed that the record of craters exposed at the Earth's surface is almost complete for crater larger than 6 km. However about 300 impact craters smaller than 6 km in diameter are still expected to be found. The minority of the discovered impact craters is pristine; most are heavily eroded and hardly to detect, or they are covered by younger deposits. For the confirmation of terrestrial craters unequivocal evidence in the form of shock metamorphic effects in minerals and rocks has to be documented. Among those effects are shatter cones, high-pressure polymorphs (e.g. coesite, stishovite, diamond), diaplectic glasses, various forms of planar features in quartz and other minerals (planar fractures, feather features, planar deformation features).

In the outer solar system the icy satellites contain abundant impact craters. Heavily cratered surfaces like that of the Saturnian satellite Rhea indicate stable and long-lasting surface conditions. Other icy bodies, such as Europa show thermal evolutions that led to a repeated recycling of an ice-crust. This reduces the population of impact craters and those that are present may be strongly deformed or modified. A surprisingly low impact crater density was recently found on the dwarf planet Pluto and its satellite Charon by the near-flyby of the New Horizons spacecraft indicating a high resurfacing rate on these planetary bodies.

Excursion 7.1 (Impact Craters as Chronometers)
Impact craters are a fascinating study topic and provide many insights in the evolution of planetary bodies and their compositional properties. Impact craters can do even more: they are the most important tool for planetary remote-sensing geologists to do what they are trained to do: mapping lithologic units and attaching ages in order to build a stratigraphic record.

Geologists obtain ages on two different ages scales: a *relative* age and an *absolute* age. A relative age can be derived from geologic mapping using classic geologic principles (see Chap. 2) and it tells us which unit is older and which is younger, but it does not provide the actual age. Absolute ages, however, are more difficult to extract. While many planetary samples have been brought to Earth by meteorites from places unknown, only few local

(continued)

[1] http://www.passc.net/EarthImpactDatabase/, August 2017.

Excursion 7.1 (continued)
spots on the Moon have been sampled during the Apollo and Luna programs. Laboratory measurements on the Earth provided high-precision radiometric ages for these rocks and with that few geologic units on the Moon could be associated with an age.

In the same way as Steno made very basic assumptions (see Chap. 2) it has been hypothesized, that older surfaces on the Moon show more and larger impact craters than younger surface units. In order to derive a *relative age* from remote-sensing image data, one can either use Steno's principles, or one can start counting craters and their sizes for a well-defined geologic unit and build a so-called *Crater-Size Frequency-Distribution (CSFD)*. The CSFD shows the total number of impact craters as a function of crater-diameter sizes. This distribution can be compared to measurements from other units to make a statement about which unit is older and which is younger.

In order to derive an *absolute age*, however, one needs to calibrate crater-size frequency-distributions with absolute ages based on Apollo and Luna measurements. These measurements allowed to establish a so-called *chronology model* which is characteristic for each planet as it describes the impact rate over time (see Fig. 7.1).

By fitting the CSFD of a measurement area to the planet's characteristic production function we obtain the number of impact craters per unit area for a given diameter size. This frequency of impact craters is the actual connection between the planet's production function and the chronology

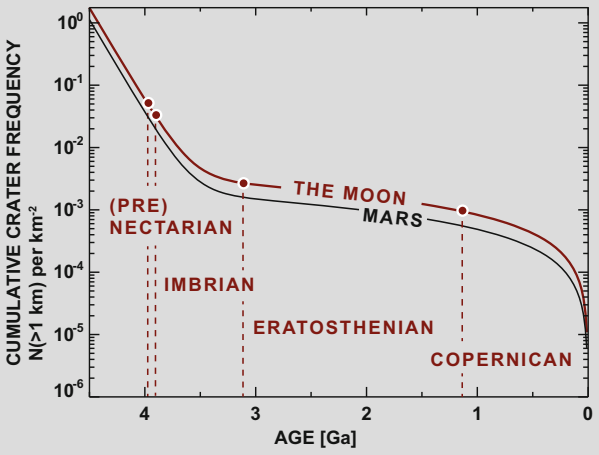

Fig. 7.1 Chronology functions for the Moon (with periods) and Mars (figure done by S. van Gasselt and A.-P. Rossi)

(continued)

Excursion 7.1 (continued)
model and allows researcher to derive absolute ages. In order to perform such measurements on other planets as well, the lunar chronology model needs to be transferred using a number of assumptions on the impact dynamics in the neighborhood of the planet and on the planet's characteristics (gravity, atmosphere, target properties and others).

7.2 The Impactor Flux

The asteroid flux has been roughly constant during the last 3.6 Ga. The earlier epoch, the so-called *Late Heavy Bombardment (LHB)*, about 3.9 Ga ago, was a phase of a much more intense bombardment.[2] During this period the Mare basins on the Moon were formed. Dynamic modeling of the solar system suggests that this phase was induced by an orbital destabilization of the four giant gas planets, which caused an injection of a great amount of main-belt asteroids and planetesimals from the outer regions to the inner part of the solar system.

Impacts in the inner solar system, in particular on the Earth-Moon system and on Mars are mainly caused by a population of objects, called Near-Earth Objects (NEO). The residence time as NEOs is limited to a few million years. It is believed that there is a steady-state orbital equilibrium between consumption of NEOs by impact processes with planetary bodies on the one hand and a replenishment of the population from the main asteroid belt on the other hand. Source regions of the NEO population are the Kirkwood gaps of the asteroid belt. These narrow orbital regions are in resonance with Jupiter or show a secular resonance with Saturn and result in a change of the eccentricity and inclination of the small bodies that are captured in these gaps. The Kirkwood gaps are constantly supplied by smaller asteroids via the so-called Yarkovsky effect. This effect is caused by an anisotropy of the thermal re-emission of the light received from the sun on the surface of a small asteroid ($>$20 km) in rotation and causes a slow drift of the asteroid into the resonance orbits.

Based on the discovery statistics and dynamic modeling there should be approximately 1000 NEOs larger than 1 km, of which almost 90% were already discovered

[2]This topic is currently matter of intense debate and different views exist, though. Measurements imply that the *Early Heavy Bombardment* during the early years of planet formation might not have simply dropped off to a near-constant rate. Instead sample ages show a potential spike in the distribution at around 3.9 Ga ago which might imply a second phase of bombardment: the *Late Heavy Bombardment (LHB)*. However, the existence of the LHB is not commonly accepted today and further sampling will be needed to show if this spike at 3.9 Ga is real or if it based on sampling errors with all its implications for current modelling the dynamic settings in the early Solar System.

by space surveys such as LINEAR or CATALINA. The number of NEOs larger than 150 m is estimated to be in the order of 30,000. Of those slightly more than 12,000 are discovered so far. The NEO website maintained by NASA JPL[3] provides daily updated information about NEO populations and close approaches.

Relatively little is known to which degree Kuiper-Belt Objects (KBO) and comets contribute to the impact crater rate in the inner solar system. However, they represent the main population of impactors for the outer solar system.

7.3 The Three Stages of Impact Cratering

Although the formation of an impact crater is a continuous and very rapid process, a subdivision into three successive stages has proven very useful to distinguish different processes of cratering.

7.3.1 Contact and Compression

The *contact and compression stage* is the first of the three stages of impact cratering and comprises the period from the initial contact of an impacting body with the ground until the compression ends. The velocity of the impacting body (projectile) depends on the gravitational attraction of the planetary body (target), the orbital speed of the projectile with respect to the target, and a possible deceleration during passage through an atmosphere. The mean collision velocity of asteroids with the Earth is approximately 17 km/s. Upon contact with the ground surface the kinetic energy of the impacting projectile is instantaneously partitioned into internal energies (deformation, heating) of the projectile and the target and the remaining kinetic energies of both projectile and target. This conversion results in the generation of shock waves that develop at the interface of projectile and target. Shock waves are characterized by an abrupt, nearly discontinuous change in pressure, temperature and density. They travel through media at a higher speed than an elastic wave and induce a flow of the media they have traversed, cause irreversible deformation within the media that is known as shock metamorphism, and heat the material. Mass, momentum, and energy are conserved during the transition from an unshocked to a shocked state, but not entropy. An equation of state relates pressure, specific volume, and the internal energy for each material.

The amplitude of the shock wave decays with distance travelled by the shock wave due to geometric spreading and consumption of energy and usually follows a power law decay curve. The degree of attenuation is material dependent, e.g. in porous rocks such as regolith or rubble piles the shock wave magnitude decays

[3] http://neo.jpl.nasa.gov/stats/.

much faster than in low-porosity rocks. Behind the shock wave, a rarefaction wave generated by reflection of the shock wave at the free surface releases the compressed material from its high-pressure state and finishes the first phase of impact cratering. The passage of the shock wave results in a net increase in temperature. The duration of the first phase depends on the diameter of the projectile and the shock wave velocity and lasts only for several tenths of a second, even for 10 km projectiles.

7.3.2 Excavation

The *excavation stage* comprises the opening of the cavity and includes the expansion and dissipation of the shock wave. As the shock wave propagates hemispherically from the initial contact area, the engulfed material is compressed and accelerated. Upon unloading the material is not completely decelerated and a residual velocity remains. This residual velocity plays a crucial role in impact crater excavation and is the most important aspect that distinguishes hypervelocity impacts from low-velocity impacts.

Near the point of impact, streamlines of the excavation flow are directed downward and outward, away from the point of impact. Particle trajectories gradually bend upward and outward caused by the reflection of the shock wave off the free surface. The flow field leads to the opening of a cavity that has a parabolic shape in cross-section (Fig. 7.2a).

Ejecta do not include material excavated from the full depth of the cavity. Target material from the deeper parts is displaced downward into the ground and does not leave the cavity. It is only the material from the upper one-third to half that is entrained in the excavation process and is deposited as allochthonous ejecta. The cavity stops growing when the remaining kinetic energy is insufficient to displace the target against its own weight (gravity-dominated cratering) or to overcome the cohesive strength of the target material (strength-dominated cratering). The resulting unstable cavity at the end of the excavation stage is called the transient cavity, and its diameter can range between a few projectile diameters to several tens of projectile diameters, depending on projectile velocity and size, and gravity. Experimental observation of transient craters in particulate targets yields depth-diameter ratios of roughly 1/3 (Fig. 7.2a).

7.3.3 Modification

In the *modification stage*, the direction of flow is reversed with respect to the excavation stage. The modification stage begins when the excavation flow comes to a halt and the transient cavity reaches its largest horizontal extent at the level of

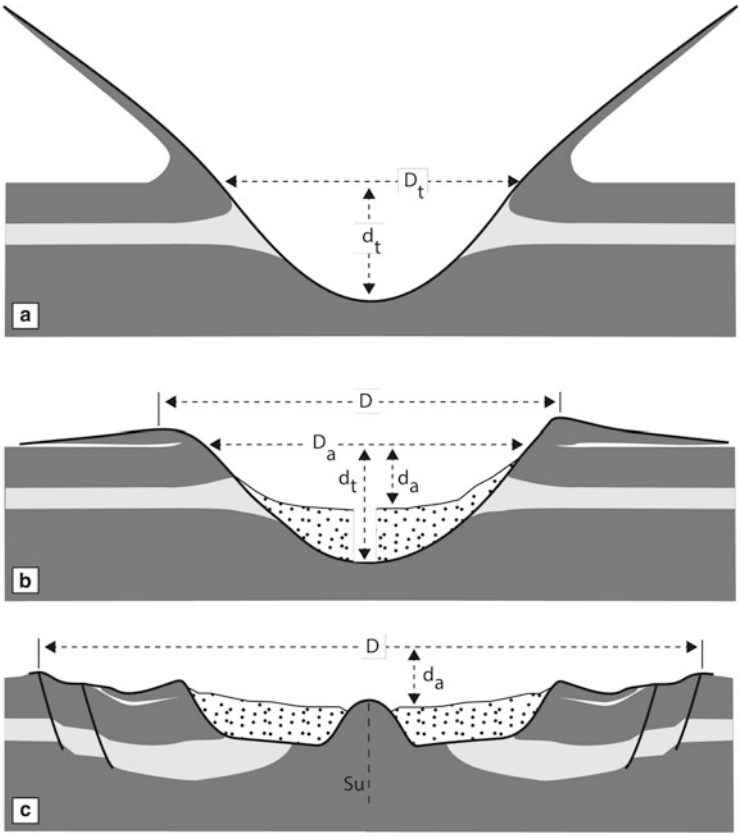

Fig. 7.2 Crater cavities and morphometric parameters; (**a**) transient crater cavity formed at the end of the excavation stage. The transient depth-to-diameter ratio d_t/D_t is approximately 0.33; (**b**) simple craters keep the principle outline of the transient cavity, in contrast to (**c**) complex craters. D is the crater diameter, D_t refers to the transient cavity diameter, d_t is the transient cavity depth, d_a is the apparent crater depth and *Su* denotes the amount of structural uplift

the target surface. It is important to mention here that the excavation flow does not stop simultaneously in all parts of the transient cavity before reversing its direction.

Gravity is the principal force that drives the collapse of the transient cavity. Depending on the degree of modification, the final crater is classified into either a simple or complex morphology (Fig. 7.2b, c). Simple and complex impact craters have fundamental differences in morphology and structure; both being a function of size. Simple craters generally have a bowl-shaped morphology (Figs. 7.2b and 7.3a), while complex craters have terraced rims and an uplifted central portion of the crater floor (Figs. 7.2c and 7.3b). With increasing size, complex craters are further

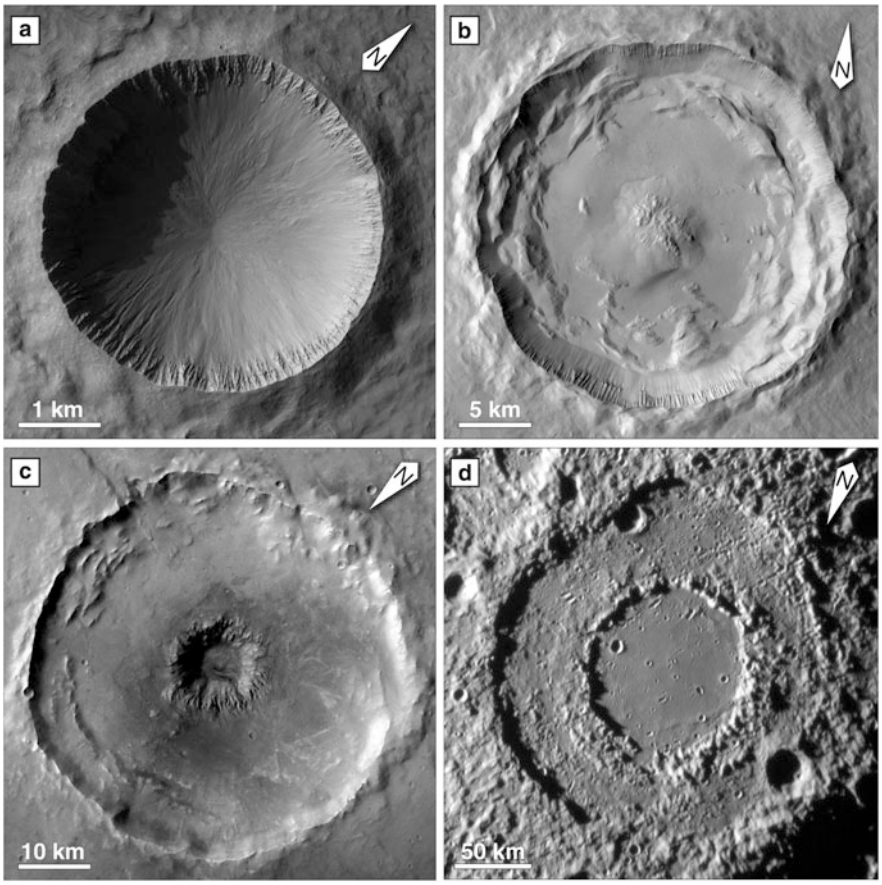

Fig. 7.3 Examples of impact crater morphologies in the solar system. (**a**) Bowl-shaped *simple crater* on Mars; (**b**) complex crater on Mars with central peak (*central-peak crater*), flat crater floor and terraced crater rim; (**c**) *central pit crater* on Mars; (**d**) Michelangelo *peak-ring crater* on Mercury. Source: (**a**) MRO/HiRISE. (**b**) MRO/CTX. (**c**) MEx/HRSC. (**d**) Messenger MDIS

subdivided into central peak (Fig. 7.3b), peak-ring craters (Figs. 7.3d, 7.4a, and 7.5), and multi-ring basins (Fig. 7.4b). Simple and complex crater morphologies have distinctly different depth-diameter ratios (Fig. 7.6a). Comprehensive mass movements make complex craters much broader than simple craters. The comparative planetological analysis of the simple-to-complex transition of impact craters from various planetary bodies shows that the transition diameter between simple and complex morphologies is inversely proportional to the surface gravity (Fig. 7.6b). This indicates that gravity is the main driving force for crater modification. However, the transition is also controlled by the strength.

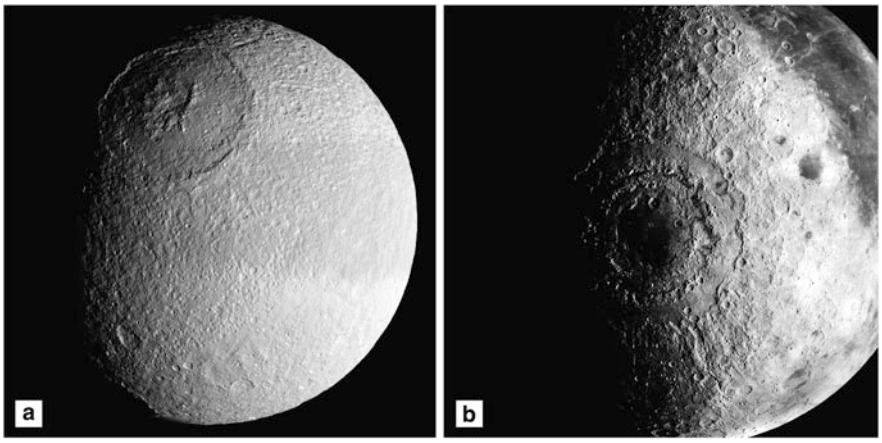

Fig. 7.4 Examples of impact crater morphologies in the solar system. (**a**) Large *complex pit/peak-ring crater* Odysseus on the icy moon Tethys that is near the threshold size for disruption; (**b**) The *multi-ring basin* Mare Orientale on the Moon. Source: (**a**) NASA/JPL/SSI. (**b**) Lunar Orbiter IV

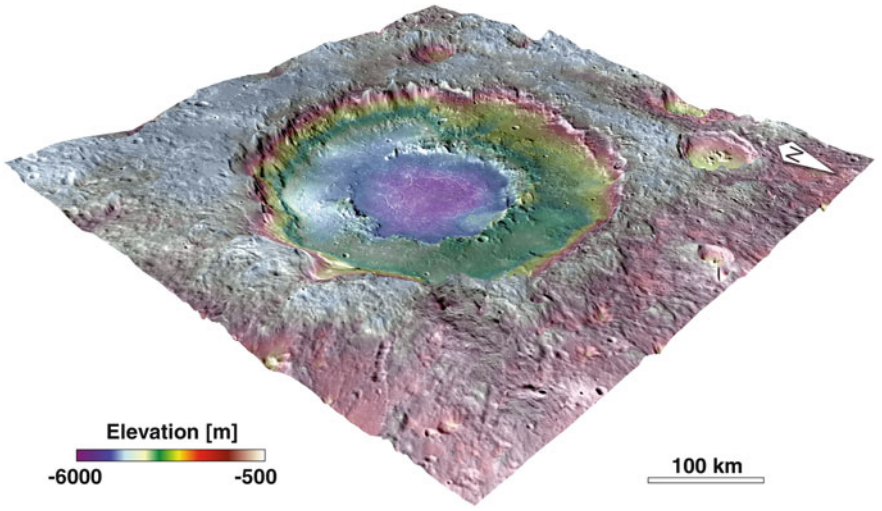

Fig. 7.5 Rachmaninoff crater on Mercury shows a complex morphology including a ring of peaks in the center part and smooth plains inside the peak ring with a set of concentric troughs. Source: Messenger MLA/MDIS

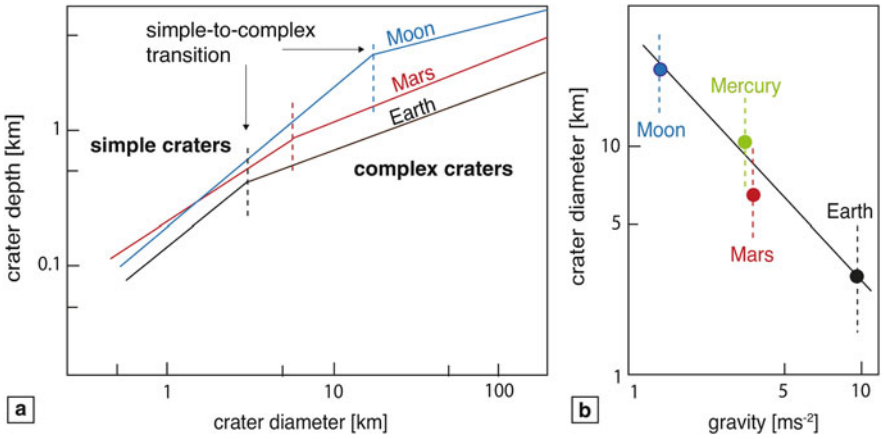

Fig. 7.6 (**a**) The depth-diameter ratio displays a characteristic kink that marks the simple-to-complex transition; (**b**) the simple-to-complex transition diameter is inversely proportional to gravity indicating that the flow is gravitationally induced

7.4 The Morphology of Impact Craters

7.4.1 Simple Craters

The principal shape of a simple crater consists of a bowl-shaped depression and a raised crater rim and is thus similar in shape to the transient cavity (Figs. 7.2b and 7.3a). Modification is mainly restricted to mass movements along the steep transient cavity rim and the presence of brecciated material that fills the cavity. The visible floor of the crater is underlain by a lens of allochthonous target-rock breccias. Slumps initiate along the steepest parts of the crater walls near the slope top end and lead to an increase in cavity diameter by 10–20%. Incision by gullies is typical for the upper part of the cavity slope whereas fans develop downslope. Prime examples of young, well-preserved and well-documented simple impact craters on Earth are the 1.2 km diameter Barringer Crater (AZ, USA) or the 3.4 km diameter Pingaluit crater (Canada) (Fig. 7.8b).

7.4.2 Central-Peak Craters

A central peak forms when buoyancy governs the strength of the target. This happens when the transient cavity exceeds a certain threshold size. Uplift starts at the deepest point of the cavity in the center, where the upward direct buoyancy force is strongest (Figs. 7.2c and 7.3b) and spreads out sidewise. Eventually the uplift of the central crater floor occurs simultaneously with a downward and inward motion of rock masses from the cavity walls so that a complex material flow field develops.

The result is a relatively flat crater with a central peak surrounded by a ring syncline that is filled with various types of breccias and melt lenses to form an annular flat plain. The rim area is sculptured by circumferential terraces that are divided by crater inwardly dipping scarps.

The convergent flow during crater modification leads to a radially compressive stress field and is associated with intense deformation that culminates within the central uplift. The deformation inventory of central uplifts is extremely complex and comprises tight folding, faulting, and complete brecciation. It could be unraveled only for a few terrestrial and Martian impact craters with a stratified target (Fig. 7.7c, d), where the interior of the central uplift is exposed due to later erosion.

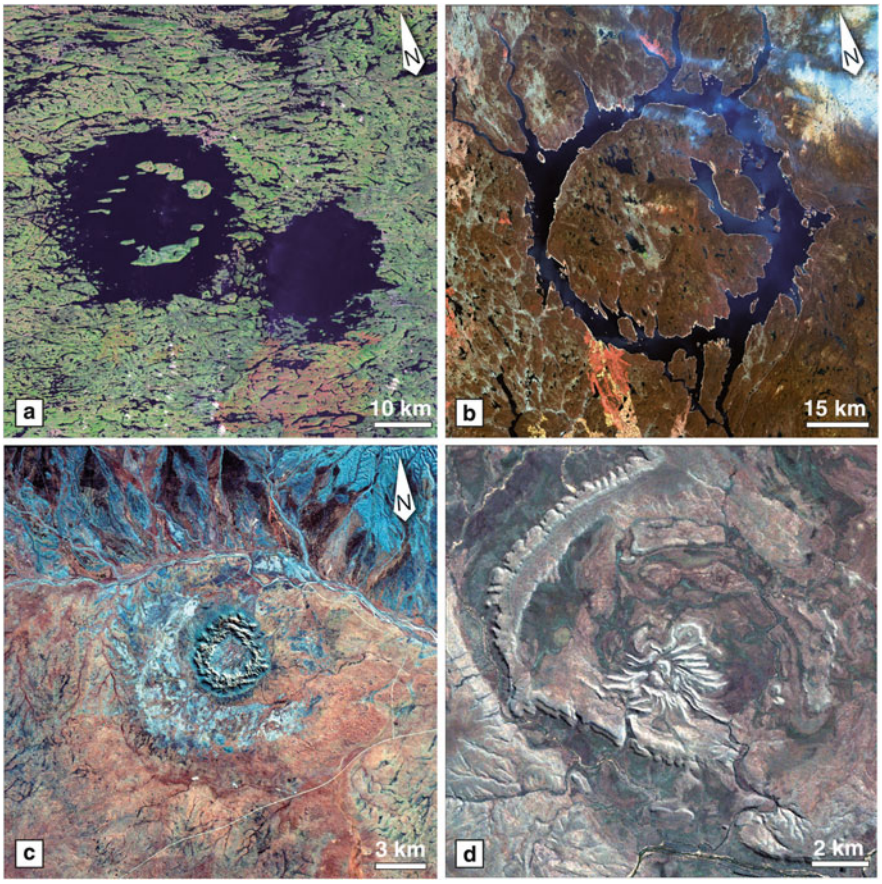

Fig. 7.7 Impact craters from the Earth; (**a**) the simultaneously formed Clearwater lakes twin impact craters (36 and 26 km diameter); (**b**) the 100 km Manicouagan crater, Canada, contains a massive impact melt rock sheet in the center; (**c**) 24 km Gosses Bluff crater, Australia, a classical central peak crater; (**d**) The central uplift of the 14 km Spider crater is composed of a stack of thrusts that indicate the impact direction from NNW to SSE. Source: (**a**–**d**) USGS/NASA Landsat

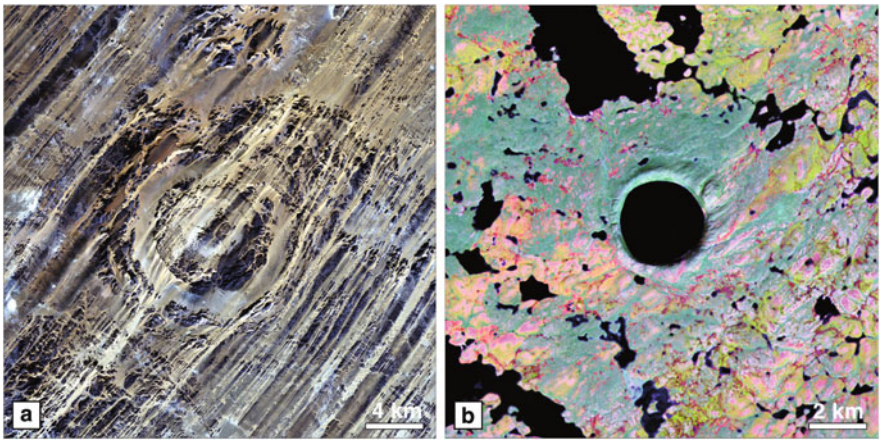

Fig. 7.8 Impact craters from the Earth; (**a**) the deeply eroded Aorounga Crater (Chad, 16 km diameter) with central peak structure and impressive yardangs; (**b**) the 3.4 km diameter Pingaluit crater (Canada) is a pristine simple crater on Earth. Source: (**a**–**b**) USGS/NASA Landsat

In pristine complex craters the diameter of the central uplift is about 1/4th of the final crater diameter. However, field analysis of terrestrial impact craters eroded to different levels and numerical modeling have proven that the centrally uplifted area becomes broader with depth while the observable stratigraphic uplift decreases. Simultaneously, the diameter of the crater shrinks with depth because the crater rim normal faults that delineate the crater usually dip inward. Both circumstances cause the increase of the ratio of the central uplift diameter to the apparent crater diameter with increasing depth of erosion. Consequently, the central uplift is by far the dominant structural feature of deeply eroded craters; the moat and crater rim become faint features at depth (Fig. 7.8a).

Central pit craters are a special type of complex crater morphologies that host a central depression at the crater center either on top of a central uplift (summit pit) or directly at the crater floor (floor pits) (Fig. 7.3c). Central pit craters are relatively abundant on ice-rich planetary bodies such as Mars, Ganymede and Callisto. A few summit pits are reported from Mercury and the Moon. Pit formation may occur via the collapse of surface material into a void left by the gradual release of impact-induced volatiles or the drainage of impact melt into sub-crater fractures.

7.4.3 Peak-Ring Craters

At a certain threshold diameter, central uplifts themselves become gravitationally unstable immediately after their formation and start to collapse under their own weight to form a morphological ring of peaks (Figs. 7.3d, 7.4a, and 7.5). This particularly occurs when the central uplift rises for a short time high above the

target surface (overshooting). The subsequent collapse is facilitated by an outwardly directed flow. These masses collide and overthrust with the inwardly moving blocks of the surrounding annular trough. The hummocky morphological appearance of fresh peak-ring structures on the Moon or on Mercury (Figs. 7.3d and 7.5) suggests that the intensity of collapse varies sector-wise. The onset of peak-ring formation also relates inversely to gravity of the particular planetary body. On the Moon peak-ring craters develop at sizes larger than about 140 km, whereas on Mars the smallest peak-ring craters are only 45 km in diameter. Examples for peak-ring craters on Earth are for instance, the Manicouagan crater in Canada (Fig. 7.7b), or the 65 Ma Chicxulub impact crater in Mexico. This crater was buried beneath younger Paleogene strata. Its structure could be deciphered only by means of reflection seismic surveys and deep drilling campaigns. The rugged peak ring of Chicxulub has a diameter of roughly 40 km radius and stands several hundred meters above the otherwise relatively flat crater basin floor. Inward dipping reflectors beneath the peak ring are interpreted to result from an outward thrusting of the central dome onto the annular trough.

7.4.4 Multi-Ring Craters

The largest impact structures known on planetary surfaces are so-called multi-ring basins. They can reach diameters of several thousand kilometers (e.g., the 3800 km Valhalla multi-ring basin on Callisto). Multi-ring basins are known from Mercury (e.g., Caloris Basin basin, 1300 km in diameter), the Moon (e.g., Orientale, 900 km in diameter, Fig. 7.4b), and Mars (e.g., Argyre, 1800 km in diameter) and icy satellites Ganymede and Callisto. Precise measurements of the diameters of the multi-ring basins are often scattering because it remains controversial which of the multitude of concentric and irregular rings is considered as the crater diameter.

The remotely visible circumferential rings are the traces of faults and associated escarpments. Several hypotheses tried to explain the formation of multi-ring basins. The most accepted and reliable theory correlates multi-ring formation with the thickness of the lithosphere. Multiple ring structures develop when the depth of the transient cavity is comparable to the thickness of the elastic lithosphere. For a given impact energy the number of rings primarily depends on the thickness and strength of the lithosphere and the viscosity of the underlying asthenosphere. For a very thin, weak lithosphere and underlying asthenosphere of sufficiently low viscosity (almost a liquid), the basin formation is followed by multiple oscillations of the cavity and outward propagation of gravity waves which disrupt the entire lithosphere (ripple ring basins).

For higher asthenosphere viscosities, the asthenosphere will flood the cavity but dampen the propagating gravity waves. Still, brittle concentric ring fractures will crosscut the entire thickness of the lithosphere. Thicker lithospheres restrict the number of rings that can form. Ring formation is suppressed if the lithosphere is too thick. Multi-ring basins on the Moon are associated with a bulls-eye central

positive gravity anomaly (mascon), surrounded by a collar of negative gravity anomaly. This in turn is surrounded by an outer positive anomaly, as revealed by high-resolution gravity data from the GRAIL spacecraft. This pattern arises as the natural consequence of impact crater excavation and collapse followed by isostatic adjustment and flexure during cooling and contraction of a voluminous melt pool. The primary parameters that control the observed gravity signatures of mascon basins are the diameter of the impactor, the lunar thermal gradient at the time of impact, the crustal thickness, and the mare fill.

7.5 Ejecta Facies

The excavation of rock debris during impact events and their distribution in the craters surrounding are strongly influenced by the impact energy and angle, the rheology and volatile content of the affected crust and the presence and density of an atmosphere.

7.5.1 Emplacement Under Dry Vacuum

On airless planetary bodies the debris ejected from the growing crater follow ballistic trajectories that depend on the action of gravity. The innermost ejecta are launched first, become ejected to high altitudes, and travel fastest. Ejecta originating farther from the center are launched later, move more slowly and become deposited in the ejecta blanket nearer to the crater rim. This combination leads to the formation of an ejecta curtain that has the shape of an inverted cone. The cone expands outwards and the ejecta debris of this cone are deposited on the target surface at increasing distances from the crater. Upon forceful landing of the ejecta debris, secondary craters may develop and the excavated material from these secondary craters becomes entrained in the ejecta blanket of the primary crater. With increasing distance from the crater center the ejecta blanket contains more and more locally derived debris; a process that is known as *ballistic sedimentation* which is typical for lunar (Fig. 7.9a) and Mercurian craters. In addition to secondary cratering, near-surface material is also admixed to the ejecta by scouring and dragging from the surface. The uppermost target layers are prepared to be picked up by the forcefully outward propagating blanket as they were previously slightly lifted by *spallation*, when the shock wave reflected at the free surface.

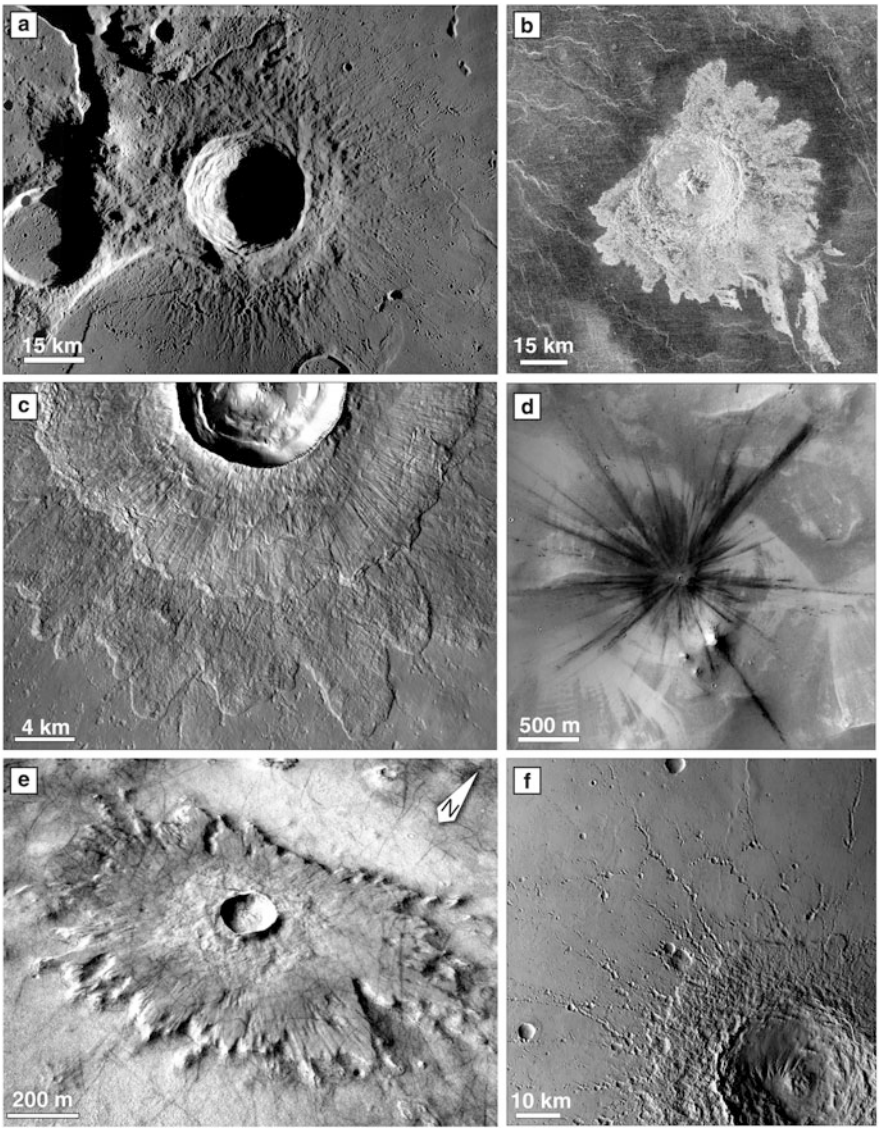

Fig. 7.9 Examples of ejecta morphologies; (**a**) the lunar Aristarchus crater as an example of *ballistic ejecta* deposition without atmospheric effects; (**b**) the Aurelia *lobate ejecta crater* on Venus with rough inner ejecta deposits with blocky material, surrounded by a lobate smoother outer layer and long flow features that extend beyond the ejecta blanket; (**c**) the Martian *double-layer-ejecta (DLE) crater* Steinheim shows two distinct ejecta layers with ramparts formed in a volatile-rich target, see also Fig. 7.10 for a perspective view; (**d**) *ray craters* possess radial crater rays that extend far beyond the continuous ejecta blanket (Mars); (**e**) *pedestal craters* are characterized by ejecta sitting above the surrounding terrain and thereby forming a raised platform; (**f**) *secondary craters* are impact craters formed by the ejecta that was thrown out of a larger crater forming clusters or radial crater trains (Mars). Source: (**a**) LRO/LROC. (**b**) NASA/JPL. (**c**, **e**, and **f**) MRO/CTX. (**d**) MRO/HiRISE

7.5.2 The Effect of Atmospheres and Target Volatiles

In the presence of an atmosphere the hot central impact plume may develop into a convective column that rises to high altitude until buoyancy ceases at a neutral density level so that the plume starts to spread out laterally. The ejecta curtain will be modified by a system of ring-like turbulences (vortex ring) in the presence of an atmosphere. Depending on the density of the atmosphere these turbulences entrain, transport and eventually deposit ejecta of various sizes into a layered pattern. The trajectories of particles involved in the ejecta curtain are thus depend not only on the action of gravity, but also on friction from aerodynamic drag and the presence of turbulences. Venus shows unique ejecta morphologies due to its thick atmosphere of about 92 bar and high surface temperatures with an average of 462 °C (Fig. 7.9b). Relatively rough inner-ejecta deposits with blocky material, likely emplaced ballistically, are surrounded by a lobate smoother outer layer. Some Venusian craters possess additionally long flow features that extend beyond the ejecta blanket and are generally interpreted as extremely fluid impact melt.

The ejecta blankets of impact craters formed in the presence of an atmosphere and volatile-rich targets, such as on Mars, possess characteristic morphologies. Such craters are typically surrounded by layered ejecta deposits that terminate in ramparts (Figs. 7.9c and 7.10). The ejecta blankets are further subdivided as single-layer ejecta (SLE), double-layer ejecta (DLE), and multiple-layer ejecta (MLE)

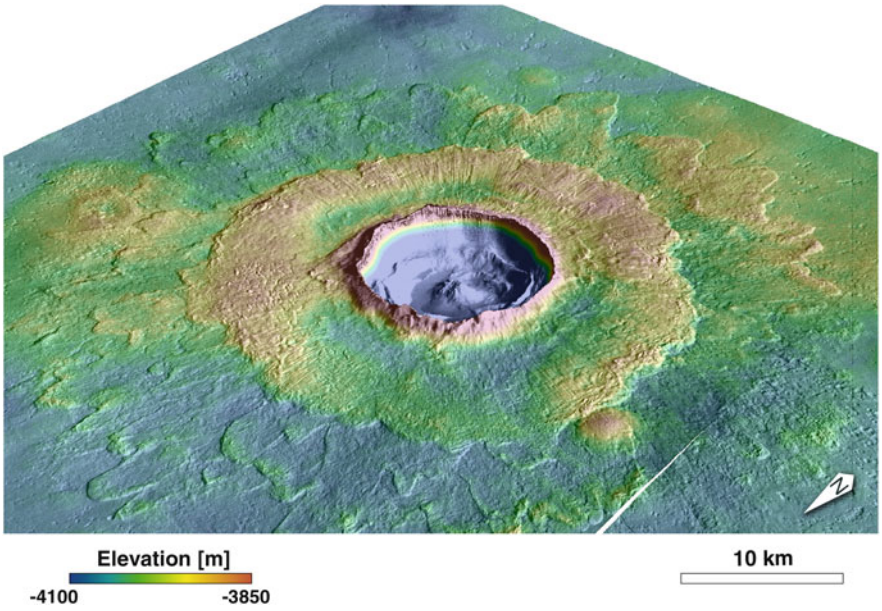

Fig. 7.10 The Martian *double-layer-ejecta (DLE) crater* Steinheim shows two distinct ejecta layers with ramparts formed in a volatile-rich target (MGS/MOLA, MRO/CTX), see also Fig. 7.9c. Source: MGS/MOLA, MRO/CTX

craters. Several mechanisms have been proposed to explain the layering and rampart formation process of Martian impact craters:

1. the interaction of ejecta with ring vortices in the atmosphere as described above,
2. the interaction of ballistic ejecta with a vapor plume formed by an impact into a volatile-rich target, causing a ground-hugging flow,
3. a combination of both,
4. formation of the layered structures by dry granular flow, and
5. formation as a landslide.

The existence of rampart craters on Ganymede indicates that volatiles in the subsurface are likely the dominant factor as there is no noteworthy atmosphere present. New results of the morphology of the ejecta blanket of the terrestrial Ries impact crater show striking morphological similarities to double-layer ejecta craters.

Another special morphology of Martian impact craters are pedestal craters (Fig. 7.9e), excess-ejecta craters, and perched craters; the latter are larger than pedestal craters. All are characterized by ejecta sitting above the surrounding terrain and thereby forming a raised platform. The ejecta formed an erosion-resistant layer, thus shielding the underlying rocks from erosion. Some pedestals stand more than hundred meters above the surrounding area and indicate how the planetary landscape evolved since the formation of the crater. It is believed that pedestal craters and related morphologies form in ice or ice-rich layers.

7.5.3 Rayed Craters and Secondary Craters

Radial crater rays (Fig. 7.9d) can extend several hundreds to even thousand kilometers across a planetary surface and are often accompanied by small secondary craters and trains of secondaries formed by larger chunks of ejecta. Such rayed craters were first identified on the Moon, where they appear brighter than the surrounding. The rays of the young crater Tycho can be observed with naked eye from the Earth. Rayed craters also occur on Mercury. Originally it was thought that such streaks exist only on planets or moons lacking an atmosphere, but more recently they have been found on Mars in infrared images taken from orbit by Mars Odyssey's thermal imager.

The high albedo of rays on the Moon is either due to a thin layer of freshly exposed regolith material or it results from a compositional difference of the fine-grained ejecta fragments covering in a ray-like manner the surrounding regolith. Space weathering from exposure to cosmic rays and micrometeorites causes a steady reduction of the difference between the ejecta's albedo and that of the underlying material. Hence the presence of rays is an indicator for a relatively young age of a crater. But the albedo also depends on the portion of iron oxide (FeO), which can retain its lighter appearance for longer periods. Low portions of FeO result in brighter materials. It is suggested that rays result from an early high-velocity ejection process induced by spallation near the point of impact. A release of the stored elastic strain in ejected fragments during the ballistic transport with

subsequent fragmentation in flight may explain why rays or trains of secondary crater develop.

Secondary craters are impact craters formed by the ejecta that was thrown out of a larger crater. Thus secondary craters surround the substantially larger primary crater, sometimes forming clusters or radial crater chains (Fig. 7.9f). Such secondary crater trains are not to be confused with crater chains (catena), where the impactor is tidally disrupted into a string of smaller objects following roughly the same orbit and forming a chain of impact craters (Fig. 7.11a). Comet Shoemaker-Levy 9 was

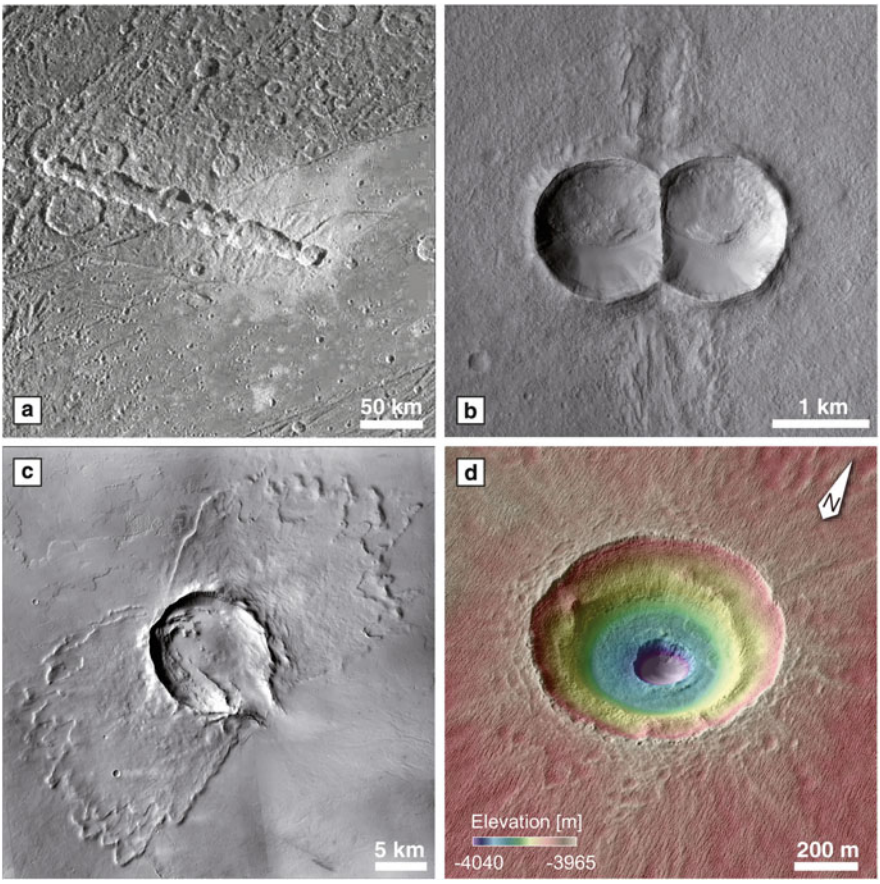

Fig. 7.11 Special crater morphologies; (**a**) where the impacting projectile is tidally disrupted into a string of smaller objects following roughly the same orbit, an *impact crater chain* can be formed (Ganymede, Enki Catena); (**b**) a *double impact crater* can be formed when a binary asteroid pair or a loosely connected bolide, that is separated into two distinct pieces prior to the impact, struck the surface (Mars). (**c**) In the case of very *oblique impacts*, the crater outline becomes elliptical and the concentration of ejecta develops a so-called *butterfly ejecta pattern* (Mars); (**d**) *nested craters* are formed due to an impact into a layered target with different mechanical properties (Mars). Source: (**a**) Galileo/SSI. (**b**, **d**) MRO/HiRISE. (**c**) MRO/CTX

disrupted in such a way prior to its impact on Jupiter and consequently, crater chains can be observed relatively frequently on Ganymede and Callisto due to the high gravity of Jupiter. If the bolide is separated into two distinct pieces prior to the impact or a binary asteroid pair struck the surface, a double impact crater can be formed (Fig. 7.11b). Secondary craters form at lower impact velocities and always result in morphologically simple crater pits. The ejecta of a secondary crater can interfere with the ejecta curtain of the primary crater. This creates v-shaped concentrations of ejecta in the surrounding of the secondary crater, towards the primary crater. Secondary craters received attention as a possible source of error for chronological studies because they may seriously affect crater counting results.

7.6 Oblique Impacts

The angle of incidence at which asteroids or comets encounter the surface of a planetary body follows a Gaussian probability distribution with vertical and horizontal collisions being extremely unlikely and the highest probability existing for 45° regardless of the planet's gravitational field. In principle, there are different approaches to derive the angle and azimuth of an oblique impact: (1) the crater morphology, (2) the distribution of the ejecta and (3) the internal structure of the crater's sub-surface. The following sections will show that the crater morphology is rather insensitive to the impact obliquity. Methods (2) and (3) are complementary.

7.6.1 Crater Outline

The shape and morphology of the crater rim is largely insensitive to the impact trajectory and remains circular with the exception of highly oblique impacts (Fig. 7.11c). Remote-sensing studies of impact craters on Mars, Venus and the Moon reveal an elliptical threshold angle of approximately 12° and an about 4% population of elliptical craters with an aspect ratio of 1.2. On Earth there is only one confirmed crater in Australia (Matt Wilson) that shows an elliptical outline. Other morphological criteria have been cited as diagnostic of oblique impacts, such as a depressed crater rim with a steepened inner slope up-range, or a larger central uplift diameter than normal relative to crater diameter. However, none of these factors has been studied statistically. An offset of the central peak from the center was statistically tested for Venusian and Lunar mare craters. While there was no unequivocal statistical result, an off-set either in uprange or in downrange direction seems to be preferred.

7.6.2 *Ejecta Distribution*

The ejecta distribution starts to become sensitive if the impact angle is less than 45°, measured from the surface. A preferential concentration of ejecta occurs downrange and a *forbidden* ejecta zone develops uprange (Fig. 7.9b). At extremely shallow angles, *butterfly* ejecta blankets form (Fig. 7.11c), where the majority of ejecta is distributed cross range.

7.6.3 *Central-Uplift Structure*

Craters that undergo erosion, first lose their ejecta blanket. The over-whelming majority of terrestrial crater fall in this category, but also many Martian craters. In such cases, the crater's sub-surface, in particular that of the central uplift, becomes exposed and provides important information of the cratering kinematics. The subsurface structure of several terrestrial impact craters as well as Martian craters show a preferred strike and dip orientation of strata and fault planes that systematically deviate from axial symmetry and imply a preferred transport direction during the crater modification process from up range to down range due to the remnant horizontal momentum in an oblique impact. Such deviations occur, for instance, at Upheaval Dome (USA), Spider (Australia, Fig. 7.7d), and Gosses Bluff (Australia) (Fig. 7.7c). The most important structural features indicative for an oblique incidence are: (1) dominance of strata strike perpendicular to the impact direction (Fig. 7.7d), (2) dominance of a thrust direction within the central uplift in downrange direction leading to a stacking sequence, (3) bilateral symmetry of the central uplift structure, with an axis of symmetry corresponding to the impact trajectory, (4) occurrence of anticlines and synclines parallel to the symmetry axis, (5) faults with strike-slip component trend parallel to the impact trajectory.

7.7 The Influence of the Target

Craters formed in layered sedimentary rocks or in ice are highly susceptible to crater collapse; whereas, craters in dense crystalline targets are relatively resistant to collapse. The standard model of crater formation applies to impact craters formed in an isotropic, uniform target, although vertical and horizontal heterogeneities of the target body are typically for a lot of planetary bodies. The so-called *nested craters* are an impressive example of craters formed by an impact into a layered target with different mechanical properties forming a set of terraces which gives it the appearance of a bullseye (Fig. 7.11d).

A large part of the Earth's surface is covered by sedimentary rocks or water, leading to a variety of different crater morphologies. For example, the terrestrial

Ries crater (26 km diameter) shows a topographic crystalline ring instead of a central peak, which would be predicted by the standard model for a similar-sized crater, due to an impact into a two-layered target of about 650 m sediments underlain by crystalline basement rocks. Vertical heterogeneities or layering of the target surface can be found in a broad range within the solar system, mostly due to variations in composition and/or temperature: e.g. (1) loose, fragmental regolith layers that overlie consolidated lava flows in the mare regions of Moon; (2) ice (water) and rock layers or rock/ice mixtures on Mars; (3) ice layers with brittle-ductile transition zones, partially underlain by water, on icy satellites, such as Europa, Ganymede and Callisto. The rheological properties of icy surfaces are of particular interest because the crater morphologies partially deviate from those observed on the terrestrial planets. Although the morphology of simple craters on icy satellites is similar to rocky bodies, a large variety of complex crater morphologies can be found (Fig. 7.4a), especially on the Galilean moons (except for Io), such as regular complex craters, central pit craters, craters with central domes, palimpsests, impact basins and multi-ring basins.

7.8 Impact Lithologies and Target Weakening Effects

Brittle and shock deformation convert the target underneath the crater floor into a shattered and brecciated mass. Crater floor breccias are commonly monomict in composition whereas breccias from ejecta blankets are polymict. The intensity of impact deformation increases from the rim to the center. This is associated with a transition from localized brittle faulting to a more pervasive cataclasis. Thus, in the crater rim zone, large-scale displacements occur on localized fault planes, that form fault-bounded terraces. The inner crater shows an increase in brittle deformation and blocks that are smaller or more internally damaged. Recent high-resolution remote-sensing data have sufficient spatial resolution to visualize the details of impact breccias and mega-blocks on Mars or the Moon.

Melt is a common characteristic of impact structures, in particular of large impact craters. Two processes are responsible for melt generation: (1) shock-related melting and (2) friction-controlled melting. Shock melting of rock-forming minerals occurs during shock unloading of strong shock waves with pressures exceeding 45–60 GPa. The products of impact melting at terrestrial impact structures range from small glass spherules, over melt lumps within suevitic breccias to local melt ponds and thick sheets of coherent impact melt rocks. Relative to the volume of the transient cavity, the volume of impact melt rock increases with crater size. In small craters, melt volume is a tiny fraction of the transient crater volume that forms an unevenly distributed and relatively thin sheet lining the final crater floor. In contrast, large impact craters contain kilometer thick sheets of impact melt rock forming a pool in the crater center. Frictional melting of rocks resulting in the formation of *pseudotachylites* can also occur during impact cratering, but these melts occur in

veins or produce local melt injections. Their scale of occurrence range from the millimeter- to the decameter scale.

The formation of a central uplift or peak ring is inconsistent with standard strength properties of rocks and requires a mechanism that temporarily reduces the strength of rocks in a very effective way. Fracturing of the target alone is not sufficient to explain the intense collapse. A possible weakening mechanism operating during crater modification is *acoustic fluidization*. The shock wave that passes through the target rocks generates scattered seismic vibrations within the fractured rock mass beneath and surrounding the crater. These vibrations result in pressure fluctuations of the ambient overburden pressure. Like on a vibrating table during periods of low pressure, frictional resistance is diminished, leading to slip events in low-pressure zones. The time- and space-averaged effect of this process is that the rock mass behaves rheologically as a viscous fluid with a certain yield strength. Hence, this mechanism elegantly explains the distributed brittle deformation and formation of discrete blocks that are orders of magnitude smaller than the size of the transient cavity. It also explains the apparent large-scale ductile deformation, as observed in many terrestrial craters.

Though most researchers favor the concept of acoustic fluidization, other reasonable models of temporary strength-weakening exist. These regard elevated temperatures as the main driver of a strength reduction. Shock heating is relevant as an important contributor to strength softening in very large impact craters. Frictional melting during the excavation and modification flow might provide sufficient lubrication to lower the strength of block contacts during the later stages of movement. Moreover, very rapid, long-distance and high strain motions along localized shear zone under relatively low normal load may cause strain-rate weakening mechanisms that explain the temporary target weakening during crater modification.

Take-Home Messages

Impact craters are omnipresent in the solar system. Impact cratering belongs to the fundamental processes in the solar system that both form planetary bodies and subsequently shape their surfaces.

The collision rate remained roughly constant for the last 3.6 Ga but was orders of magnitudes higher during the Late Heavy Bombardment (LHB) at 3.9 Ga.

The impact process is subdivided in three subsequent stages: contact and compression, excavation, and modification. Each of these phases is governed by specific physical processes that cause different material flow fields.

Impact craters are tools to derive the age of planetary surfaces and the amount of erosion. Moreover they allow to deduce environmental conditions and help to probe planetary interiors.

Suggested Readings

Barlow, N., Boyce, J., Costard, F., Craddock, R., Garvin, J., Sakimoto, S., Kuzmin, R., Roddy, D. Soderblom, L.: Standardizing the nomenclature of Martian impact crater ejecta morphologies. J. Geophys. Res. **105**(E11), 26733–26738 (2000). doi:10.1029/2000JE001258

Bottke, W., Jedicke, R., Morbidelli, A., Petit, J.-M., Gladman, B.: Understanding the distribution of near-earth asteroids. Science **288**(5474), 2190–2194 (2000). doi:10.1126/science.288.5474.2190

Collins, G., Melosh, H. Osinski, G.: The impact cratering process. Elements **8**, 25–30 (2013). doi:10.2113/gselements.8.1.25

French, B. (ed.): Traces of Catastrophe: A Handbook of Shock-Metamorphic Effects in Terrestrial Meteorite Impact Structures. LPI Contribution 954, p. 120. Lunar and Planetary Institute, Houston (1998)

French, B., Koeberl, C.: The convincing identification of terrestrial meteorite impact structures: what works, what doesn't, and why. Earth Sci. Rev. **98**(1–2), 123–170 (2010). doi:10.1016/j.earscirev.2009.10.009

Holsapple, K.A.: The scaling of impact processes in planetary sciences. Annu. Rev. Earth Planet. Sci. **21**, 333–373 (1993). doi: 10.1146/annurev.ea.21.050193.002001

Kenkmann, T., Poelchau, M. Wulf, G.: Structural geology of impact craters. J. Struct. Geol. **62**, 156–182 (2014). doi:10.1016/j.jsg.2014.01.015

Melosh, H. (ed.): Impact Cratering: A Geological Process, p. 289. Oxford University Press, New York (1989)

Neukum, G., Ivanov, B. Hartmann, W.: Cratering records in the inner solar system in relation to the lunar reference system. Space Sci. Rev. **96**(1), 55–86 (2001). doi:10.1023/A:1011989004263

Osinski, G., Pierazzo, E. (eds.): Impact Cratering. Processes and Products, p. 316. Wiley-Blackwell, Hoboken (2012). doi:10.1002/9781118447307

Schenk, P.M.: Thickness constraints on the icy shells of the galilean satellites from a comparison of crater shapes. Nature **417**, 419–421 (2000). doi:10.1038/417419a

Schubert, G. (ed.): Treatise on Geophysics, 2nd edn., p. 5604. Elsevier, Oxford (2015)

Stöffler, D., Langenhorst, F.: Shock metamorphism of quartz in nature and experiment: I. Basic observation and theory. Meteoritics **29**(2), 155–181 (1994). doi:10.1111/j.1945-5100.1994.tb00670.x

Werner, S.C., Ivanov, B.A.: 10 – Exogenic dynamics, cratering, and surface ages. In: Schubert et al. (eds.), 10 – Physics of Terrestrial Planets and Moons, pp. 327–365. Elsevier, Amsterdam (2015). doi:10.1016/B978-0-444-53802-4.00170-6

Chapter 8
Endogenic Processes

Ernst Hauber, Daniel Mège, Thomas Platz, and Petr Brož

8.1 Introduction

Endogenic processes in geology are a function of a body's internal geodynamic activity. They comprise volcanic, tectonic, and isostatic processes, which shaped the surfaces of all terrestrial planets, the Moon, and basically all other Solar System bodies with solid surfaces that have been observed in some detail. The most recent spacecraft observations have confirmed this notion, and revealed past or present endogenic activity even on bodies where this was not previously expected (Fig. 8.1). The study of endogenic processes and their resulting landforms and landscapes puts important constraints on the internal evolution (Chap. 10) and the surface history of a geologic body (Chap. 11).

This chapter focuses on rather large-scale surface features (10^0–10^6 m). Although rovers and landers can study much smaller features, they are few in numbers and are typically very restricted in their operational range (Chap. 5), although they did provide invaluable information on the magmatic evolution of their target bodies. Another important source of information, specifically on the petrology of magmatic

E. Hauber (✉)
German Aerospace Center (DLR), Berlin, Germany
e-mail: ernst.hauber@dlr.de

D. Mège
Space Research Centre, Polish Academy of Sciences, Warsaw, Poland
e-mail: dmege@cbk.waw.pl

T. Platz
Max Planck Institute for Solar System Research, Göttingen, Germany
e-mail: platz@mps.mpg.de

P. Brož
Institute of Geophysics, Czech Academy of Science, Prague, Czech Republic
e-mail: petr.broz@ig.cas.cz

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_8

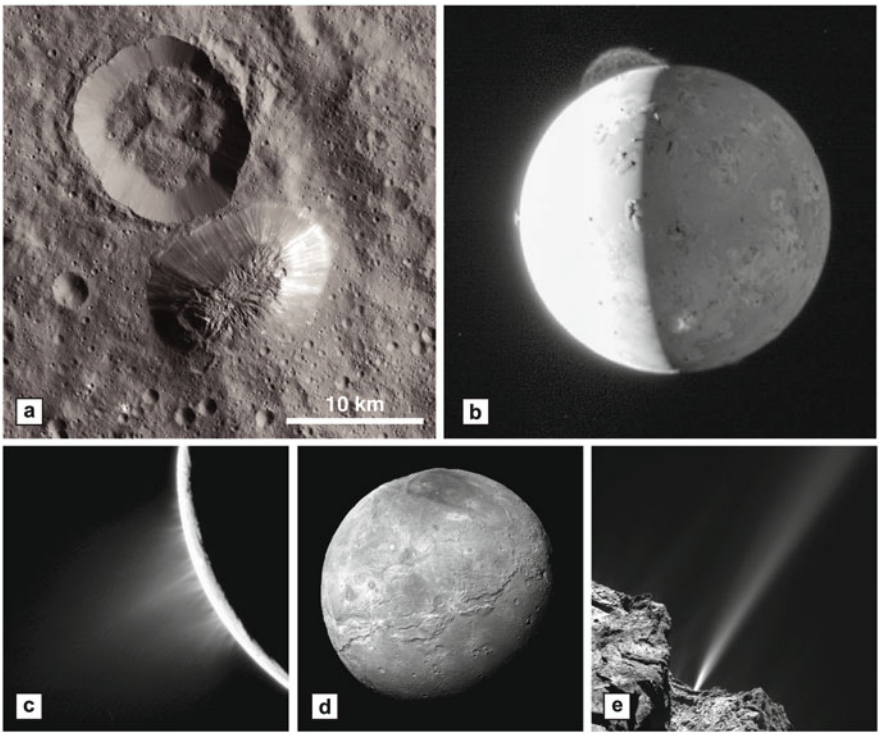

Fig. 8.1 Evidence for endogenic activity on small bodies beyond the terrestrial planets. While some of these processes were predicted on theoretical grounds (e.g., on Io) or expected from Earth-based observations (e.g., on comets), other endogenic activities came as a surprise (e.g., on Enceladus or Pluto and Charon). (**a**) A potential cryovolcanic dome on the dwarf planet Ceres. The conical edifice has a basal outline of about $\sim 10 \times 20$ km and stands about 5 km tall above its surroundings. (**b**) A plume over the active volcano, Tvashtar, on Io. The plume reaches a height of 290 km above the surface. It was imaged on 28 February, 2007 by the New Horizons spacecraft on its way to Pluto. (**c**) Jets of ice particles, water vapor and trace organic compounds emanating from the surface of Enceladus. These *ice geysers* on the ∼475 km-diameter satellite of Saturn were detected by the Cassini spacecraft. (**d**) Pluto's satellite, Charon, displays a surprisingly varied and partly young surface. The prominent tectonic fractures are evidence for stresses acting on its brittle outer shell. (**e**) A short-lived outburst from comet 67P/Churyumov-Gerasimenko. The jet is thought to have a speed of at least 10 m/s (see also Chap. 13). Source: (**a**) NASA/JPL-Caltech/-UCLA/MPS/DLR/IDA. (**b**, **d**) NASA/Johns Hopkins University APL/SRI. (**c**) NASA/JPL/SSI. (**e**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

rocks, comes from meteorites (Chap. 6), e.g., from Mars or asteroids: All such studies need to be complemented by modeling to arrive at a physical understanding of the driving forces and mechanisms.

The study of endogenic processes addresses some of the most fundamental questions in planetary geology, including the bulk composition, the history of accretion and differentiation, the heat generation and transport, and the evolution of planetary atmospheres (via outgassing) and climates (Chap. 10). Igneous intrusive

and extrusive magmatic processes operated or still operate on all terrestrial planets, possibly on larger asteroids or their parent bodies, and on some satellites in the outer Solar System, and tectonic processes left their traces on all observed objects. We first introduce the inventory of tectonic and volcanic landforms on the various planetary bodies, and present techniques how their morphology can be analyzed. We then discuss how tectonic regimes differ between the individual bodies, and how magmatism evolved on them. Given the enormous range in number and variety of endogenic processes in the Solar System, this chapter necessarily needs to limit its scope to a few major aspects. More detailed treatments of the geology of the icy satellites and minor Solar System bodies are given in Chaps. 12 and 13, respectively. Impacts can trigger significant tectonic movements and the generation of impact melt (Chap. 7).

8.2 Landforms of Endogenic Processes

8.2.1 Tectonic Landforms

The mechanical deformation of planetary lithospheres produces tectonic structures such as joints, deformation bands, faults and folds. Therefore, the study of tectonic features (*structural geology*) can help to reconstruct the history of crustal and lithospheric stresses and their causative factors, constraining models of geologic evolution and internal processes (see Chap. 10). Tectonic structures are observed on all larger Solar System bodies with a solid surface. There is no fundamental difference in the main structural elements between the Earth and most other Solar System bodies. This is not very surprising, as both the basic constituents of the deformed solids (i.e., rocks and ice) and the mechanical stresses responsible for deformation are similar throughout the Solar System. Nevertheless, the relative abundance of tectonic surface features varies widely among the planets, their satellites and asteroids. This can mostly be explained by the nature of the stresses and the tectonic style operating on different bodies. For example, strike-slip faulting is characteristic for the horizontal movement of lithospheric plates against each other and is a hallmark of plate tectonics. As Earth is the only planet where plate tectonics operate at present, it is not surprising, then, that large-scale strike-slip faulting is rare on one-plate planets such as Mars, Mercury, or the Moon.

Fractures are arguably the most common structural elements on planets and small bodies. They are typically classified into mode I fractures (opening mode: relative motion is normal to the fracture plane), mode II fractures (shear mode: motion is parallel to the fracture plane), and mode III fractures (tear mode: motion is parallel to the tip line of the fracture). Typical examples of mode I fractures are joints or tension cracks (Fig. 8.2). Although they are typically too small to be identified in remote sensing data, joints have been observed on Mars in very high-resolution images. Joints can control the preferred orientation of large-scale

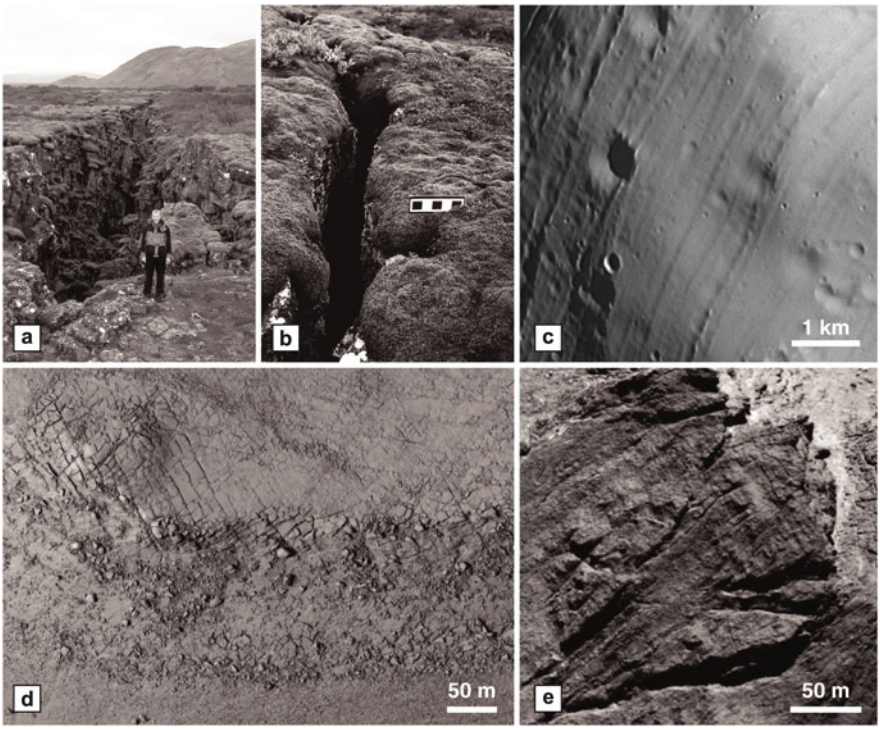

Fig. 8.2 Joints and tension cracks on Solar System bodies. (**a**) Tension fractures in the rift zone at Thingvellir, Iceland. (**b**) Close-up of tension fractures at Thingvellir: *Squares* on scale bar in are 5 × 5 cm. (**c**) The origin of grooves on the Martian satellite, Phobos, is unknown, but some studies favour a formation as tension cracks. (**d**) Joints in rocks near the Nilosyrtis Mensae region, Mars. (**e**) Fractures on comet 67P/Churyumov-Gerasimenko. Source: (**a**, **b**) E. Hauber. (**c**) ESA/-MEX/DLR/FU Berlin. (**d**) NASA/MRO/HiRISE/University of Arizona. (**e**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

erosional structures such as yardangs, hence the orientation of joint patterns can potentially be inferred from the analysis of such erosional landforms. On a much larger scale, it has also been hypothesized that some obviously tectonic fractures at the giant Valles Marineris system of linear troughs on Mars are tension fractures. Sets of linear grooves that have been detected on the Martian moon, Phobos, and on other irregular small bodies (e.g., the asteroid 21 Lutetia) may also be opening fractures, although a variety of models has been proposed to explain their origin. Most recently, new images of comet 67P/Churyumov-Gerasimenko revealed ubiquitous cracks that may have been formed as opening fractures by internal (or thermal) stresses (Chap. 13).

Mode II fractures are faults, which can further be subdivided into normal faults, reverse or thrust faults, and strike-slip faults. Normal faults are records of tensional stresses and extensional deformation, whereas reverse and thrust faults indicate compressional stresses and contractional deformation.

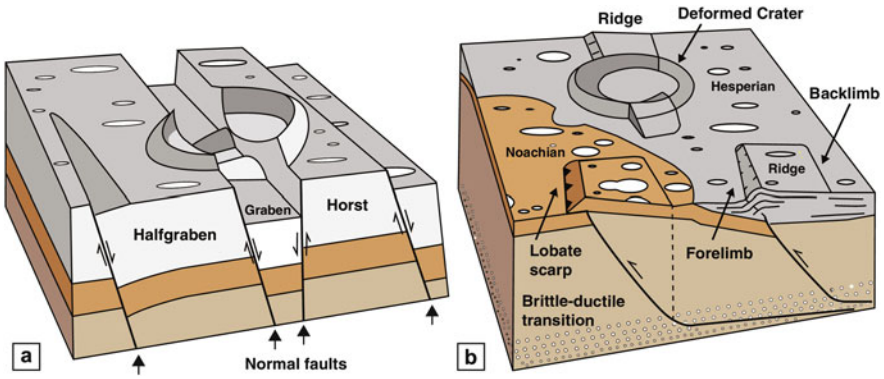

Fig. 8.3 Schematic views of major fault types on the terrestrial planets. (**a**) Extensional features. (**b**) Contractional features. Source: redrawn after Mueller and Golombek (2004)

Normal faults cut the surfaces of most, if not all planetary bodies. They are typically visible as rectilinear scarps (Fig. 8.3). Due to the lack of vegetation and the low erosion rates, fault scarps on, e.g., Mars or the Moon appear typically very pristine and can often be better studied in remote sensing data than their terrestrial counterparts. The maximum lengths of normal faults on Mars and Venus can be tens to hundreds of kilometres, although the longest faults are often composed of several linked segments. As on Earth, planetary normal faults seem to grow by segment linkage, and *en echelon* patterns and relay ramps are common. On the short end of fault lengths, they can be observed down to the resolution limit of images. Similarly, fault throw, though difficult to determine in remote sensing data, ranges from meters to kilometres. The throw is typically estimated from the vertical scarp height and an assumed fault dip. Based on the available topographic information, it seems that the displacement-length relationship of planetary normal faults is similar to that on Earth. The fault throw derived with this method can be used to constrain the extension and strain across the fault scarp. As mentioned above, these estimates all require the use of an assumed fault dip, which is typically chosen to be 60°. Based on analogy to the Earth, this assumption relies on the classical dip value for normal fault planes. It is known, however, that this value, which is based on physical principles, applies to faults at a certain depth, whereas normal faults commonly start as subvertical tension fractures and are much steeper than 60° near the surface. Unfortunately, the true dips of normal fault planes are difficult to determine from remote sensing data. Visible scarps are typically not the actual fault planes, but represent talus deposits generated by erosion and fault scarp retreat. Therefore, the measurable slope often corresponds to the angle of repose and cannot be used to constrain the true dip.

Normal faults can occur as isolated features, but more commonly they are part of a more complex structural architecture. Half-grabens (asymmetrical) and grabens (symmetrical) have been observed on the Moon, Mars, Venus, and recently also on Mercury and asteroids. They typically form sets or families, which bear records

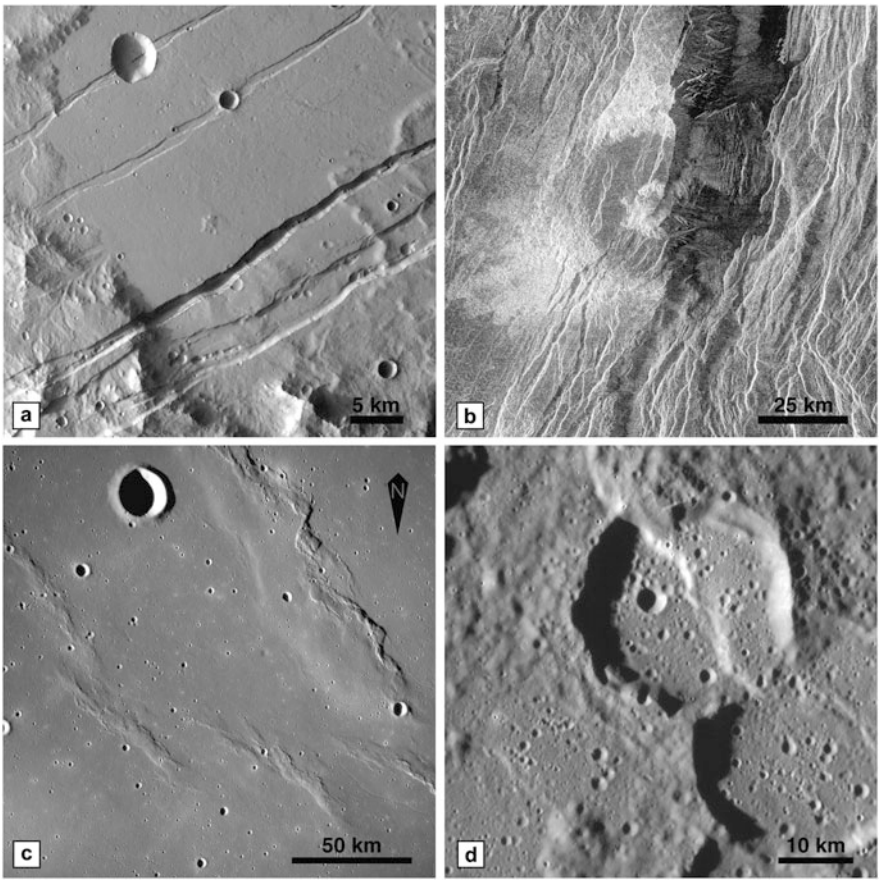

Fig. 8.4 Examples of main types of planetary faults. (**a**) Detail of a set of long and narrow grabens in the Memnonia region of Mars. North is up, illumination from left (west). (**b**) Large-scale extensional rift between Rhea and Theia Montes in Beta Regio on Venus as seen in Magellan radar data. Note rifted crater with a diameter of ∼37 km. (**c**) Lunar wrinkle ridges north of Flamsteed crater, Oceanus Procellarum. (**d**) Lobate scarp in the Rembrandt basin on Mercury. Source: (**a**) ESA/MEX/DLR/FU Berlin, HRSC orbit 4073. (**b**) NASA/JPL. (**c**) Apollo 12 image, NASA. (**d**) NASA/Johns Hopkins University APL/Carnegie Institution of Washington

of local- or regional-scale processes (Figs. 8.4 and 8.5). For example, concentric grabens along the margins of lunar maria were formed by the relatively larger subsidence in the mare interiors caused by volcanic loading (Chap. 11).

Another prominent example of planetary graben sets are the long and narrow grabens that radiate outwards from several magmatic centers in the Tharsis region on Mars, the largest known volcano-tectonic province in the Solar System.

Even more complex extensional systems resemble terrestrial continental rifts. They are only observed on larger terrestrial planets, i.e., on Mars, Venus and Earth. Venusian rifts in particular are analogous to rifts in, e.g., East Africa, and

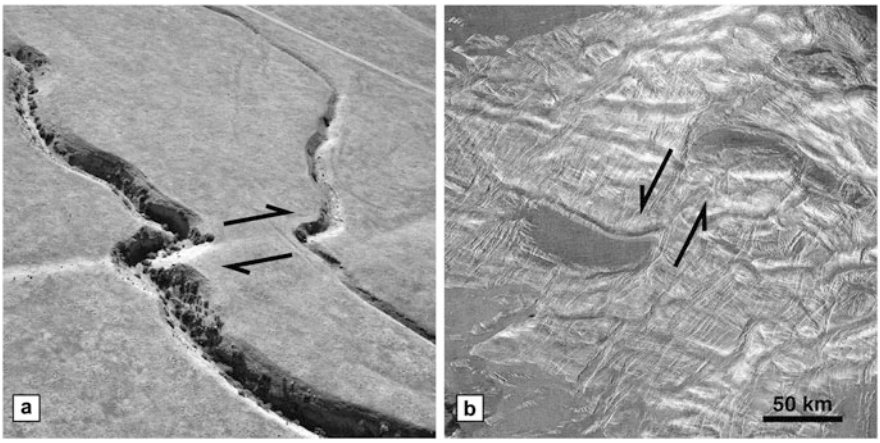

Fig. 8.5 Examples of main types of planetary faults. (**a**) Strike-slip fault as part of the San Andreas graben in California. Fault section length is ∼150 m. (**b**) Strike-slip fault in Ovda Regio on Venus. Source: (**a**) USGS photograph by David K. Lynch, Kenneth W. Hudnut and David S.P. Dearborn (2009). (**b**) NASA/JPL/Magellan

are obviously associated with major volcanic centres. Rifts on Mars and Venus display a number of properties that are characteristic for terrestrial rifts, such as rift segmentation, segment linkage at accommodation zones, and alternating rift polarity, i.e., an asymmetric cross section with a dominating master fault system at one rift margin, which changes position at accommodation zones. Extension across planetary rifts appears to be limited to a few kilometres, and strain is accordingly low, too. Terrestrial continental rifts are believed to penetrate the entire lithosphere, and indicate the beginning of continental break-up. The rifts on Mars and Venus, though similar in dimension and style, are not thought to be evidence for plate tectonics, but they are obviously major structural elements and could indicate an early stage of large-scale tensional strain localization which never developed towards a more advanced stage.

Contractional structures are widespread on all terrestrial planets. The most commonly observed class of contractional landforms are wrinkle ridges (Fig. 8.4). This descriptive term was chosen because they are characterized, in cross section, by a broad topographic rise or arch on which is superposed a smaller and steeper ridge. On the top of this ridge, an even smaller crenulation has an irregular (*wrinkled*) trace in plan-view. The ridge is typically offset to one side of the broad arch, creating an asymmetric profile with a steeper forelimb and a shallower backlimb (Fig. 8.3). The width of wrinkle ridges is variable, ranging from a few kilometres to few tens of kilometres, and the maximum height is typically a few hundred meters. Although several models have been invoked to explain wrinkle ridges, the most accepted view holds that they are thrust-propagation folds, where near-surface layered strata are folded by the motion along a blind thrust at depth (a blind thrust does not propagate to the surface and thus does not break it). Since the first wrinkle ridges studied in

detail were those on the volcanic lunar mare plains, it was once thought that they are indicative of a substrate consisting of layered basaltic lava flows. However, at least in the case of Mars, it cannot be excluded that sedimentary layers may also be folded above blind thrusts. Sometimes the plains to one side of the ridge are systematically elevated with respect to those on the other side, thus indicating the vergence of the thrust. The depth to which wrinkle ridges extend into the crust is a matter of debate. The *thin-skinned* models favour a limited depth of faulting, perhaps originating at a décollement that may correspond to a rheologic boundary layer (e.g., the brittle-ductile transition). On the other hand, a *thick-skinned* origin would involve faulting that penetrates perhaps most of the crust. Wrinkle ridges frequently display an evenly spaced pattern in plan-view, with distances of typically tens of kilometres between individual ridges. Good examples of such regular patterns can be found in the plains north and south of Valles Marineris on Mars. This spacing has been used to model the depth of the deformed layer, with shorter distances between wrinkle ridges corresponding to a thinner volume. The origin of wrinkle ridges may often be related to loading and resultant compressional stresses. For example, wrinkle ridges on the lunar mare plains are thought to result from subsidence in the mare centres due to the load of the basaltic infilling, and corresponding compressional stresses acting on the near-surface crust.

The second main class of planetary contractional features are lobate scarps. As typical for the nomenclature of planetary landforms, this term is descriptive and avoids interpretation. Lobate scarps are mostly observed on the Moon and Mercury. The characteristic landform is a topographic scarp that is lobate in plan-view (Fig. 8.3). The commonly accepted explanation is that lobate scarps are surface-breaking thrust faults. Typically, it is assumed from rock mechanics that the fault planes of lobate scarps have dips of $\sim 30°$. The topography and geometry of lobate scarps can yield clues to fault motion or lithospheric properties. The cross-sectional geometry of a lobate scarp, especially the height and the location of a trailing syncline, can be used to model the thickness of the layer that has been bent upwards by the motion along the fault plane. As this thickness is assumed to be equal to the thickness of the elastic lithosphere, this method enables estimating the heat flow at the time of faulting. The best-known example how lobate faults can be used to constrain planetary evolution is illustrated by studies of Mercury's cooling history. Mapping of the global lobate scarp population reveals their number and their individual lengths. Together with a given ratio of displacement to length, and an assumed fault dip, it is possible to calculate the total amount of cumulative crustal shortening (i.e. loss of area) that was accommodated by the lobate scarps. This surface shrinkage can then be used to estimate an associated decrease of the planet's radius.

Except on Earth (Fig. 8.5), strike-slip faulting is much less common than normal or thrust faulting on bodies with rocky surfaces: Only Venus displays some large-scale strike-slip faults (Fig. 8.5), whereas their existence on Mars is contentious. This positive correlation of increasing planetary mass with increasing importance of strike-slip faulting is not surprising, as lithospheric mobility and relative lateral motions on planets without plate tectonics (*one plate-planets*) are very limited and

basically restricted to local-scale processes. However, some small-scale strike-slip movement may have occurred in wrinkle ridged-plains on Mars and Mercury.

As compared to the ubiquitous brittle deformation, the observational evidence for ductile deformation is much sparser in the Solar System. Folding of rocks appears to be more common on Earth and Venus as compared to Mars, Mercury, and the Moon. As ductile deformation and folding occurs in deeper levels of the crust and lithosphere and is only exposed at the surface after uplift and erosion, it is typical for mountain belts, e.g., in the orogens created by plate collisions on Earth. As erosion rates are much lower on the smaller planets due to the lack of dense atmospheres, the exposure of deeper crustal levels is rare, and except on Earth, large folds have only been observed on Venus. On a smaller scale, folding due to sedimentary processes (e.g., soft sediment deformation) is apparent in high-resolution images of sedimentary rocks on Mars.

Apart from these landforms, which are all not restricted to any given planet, there are more special features that are unique to some bodies. Among the terrestrial planets, this applies especially to Venus, where a number of large-scale volcano-tectonic landforms may be related to diapirism (both upwelling and downwelling processes have been proposed). The best-known examples are the coronae on Venus, which are circular or slightly elliptical in plan-view, have diameters of up to a few hundred kilometres, and are commonly associated with concentric fracture zones and one to several volcanic centres. Other unique features on Venus are novae and arachnoids, which also reach significant dimensions, but are characterized by radial tectonic fractures. Io, Jupiter's volcanically active satellite, shows some very large mountains that seem to be uplifted along deep-seated thrust faults. The compressive forces at depth are thought to be a result of volcanic loading at the surface, which causes downward movement of the deeper crustal levels but no contraction at the surface.

8.2.2 Volcanic Landforms

Terrestrial magmatic processes produce a variety of landforms over a broad range of scales. Most of these landforms have morphological analogues on other terrestrial planets and the Moon (Fig. 8.6). The study of volcanic landforms can inform the analysis of eruption processes and, therefore, enables a better understanding of the chemical and physical characteristics of magmatic systems. Typically, volcanic landforms are classified on the basis of magma composition and the types of eruptions (e.g., effusive vs. explosive) and their products, but such classifications should also consider additional parameters such as geomorphic scale, constructional vs. erosional origin, and mono- vs. polygenesis. It is important to note that volcanic landforms are not restricted to volcanoes, but also to eruption products that may have been emplaced far from any vent (e.g., lava flow fields, ash fall deposits).

Magmatism is strongly coupled to tectonism; hence volcanic landforms are typically not randomly distributed over a planet's surface and rarely occur isolated.

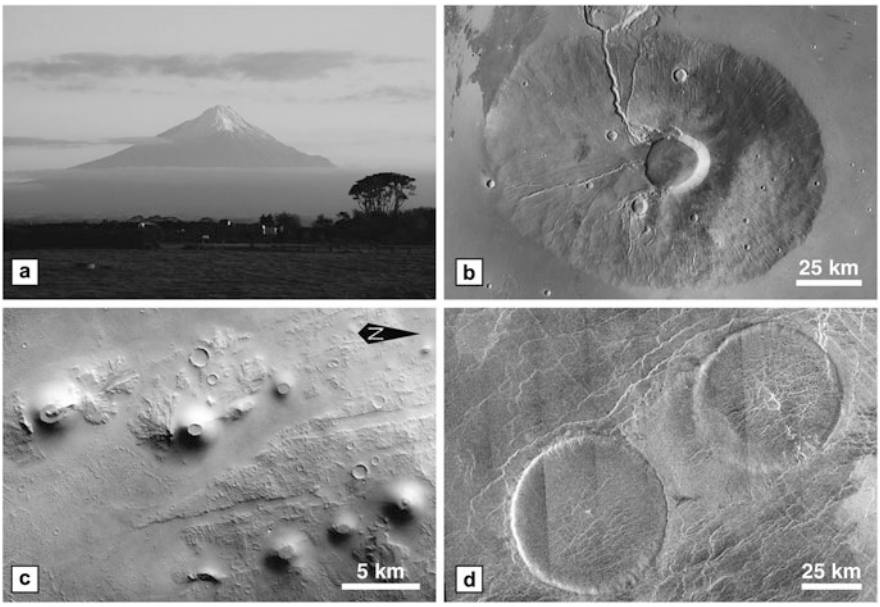

Fig. 8.6 Examples of volcanic edifices. (**a**) Mt. Taranaki (New Zealand) is an andesitic composite volcano rising from sea level to 2518 m. The summit crater hosts remnants of a lava dome. (**b**) Ceraunius Tholus located in Mars' Tharsis volcanic province is a large shield volcano. It is partially buried by the surrounding lava plains. The summit is marked by a near circular caldera. (**c**) A complex of pyroclastic cones is located in the Ulysses Fossae area north of Biblis Tholus, Mars. (**d**) Two steep-sided, flat-topped volcanic domes located in Tinatin Planitia, Venus, are shown on this Magellan radar image. They formed by extrusion of highly viscous lava. The largest dome is 62 km across. Source: (**a**) T. Platz. (**b**) ESA/MEX/DLR/FU Berlin. (**c**) NASA/MRO/CTX. (**d**) NASA/JPL/Magellan

Instead, their spatial distribution reflects the nature of endogenic processes that produce magmatism. On Earth, most magmatic activity is associated with plate tectonics, and the majority of volcanic landforms are concentrated at divergent and convergent plate boundaries. However, there are also huge volcanic and plutonic deposits and provinces that lie outside tectonic plate boundaries. They often form extended continental flood basalts and oceanic plateaus, and are collectively termed Large Igneous Provinces (LIPs). LIPs are typically characterized by effusion rates that are much higher than those observed at divergent plate margins, and are important because they enable studying lithospheric and mantle processes that are often obscured by plate tectonics. It is commonly assumed that LIPs form when the head of a mantle plume impinges on the crust and large amounts of magma are emplaced as intrusions and basaltic lava flows (but not all researchers agree on the link of intraplate volcanism to plumes). On other planetary bodies, the distribution of volcanic surface features seems to be related to narrow or broad sites of mantle upwelling. For instance, the vast volcanic plains on the Moon and Mercury are concentrated at the nearside and at a few basins and the North Polar Region,

respectively. On Mars, volcanism has been active over most, if not all of its history. Whereas older (~3.9 Ga) volcanic features are globally widespread, Amazonian volcanism (<3.4 Ga) has basically focused on two distinct large provinces, Tharsis and Elysium.

At smaller scales, volcanoes commonly tend to form clustered volcanic fields. Hundreds of individual edifices can be clustered in a single volcanic field, and the analysis of their spatial distribution can reveal structural trends that controlled their emplacement. If an individual edifice was formed during one single eruptive cycle, it is called a monogenetic volcano, and accordingly a monogenetic volcanic field consists of many monogenetic volcanoes. A typical example for a monogenetic volcano is a scoria cone, and actually scoria cones are the most common volcano type on Earth. In contrast, polygenetic volcanoes (e.g., composite volcanoes) form over several eruptive cycles.

Calderas are volcano-tectonic structures common on both Earth and Mars: they represent collapsed roofs above shallow magma chambers that were partly emptied: such landforms are rather common on Mars on various volcanic edifices, including very large ones, e.g., at the top of Olympus Mons (Fig. 8.7).

As for the Earth, planetary volcanic landforms are frequently classified according to the associated eruption type (Table 8.1). Most authors distinguish only a handful of eruption types, typically as a function of the way how the gas exits the magma. If magma ascends quietly to the surface and is not significantly disrupted by gas release, an *effusive eruption* produces lava flows. They can pile up to form mostly basaltic shield volcanoes such as on Hawaii, but more importantly in the geological past, such eruptions also contributed to the formation of huge flood basalt provinces,

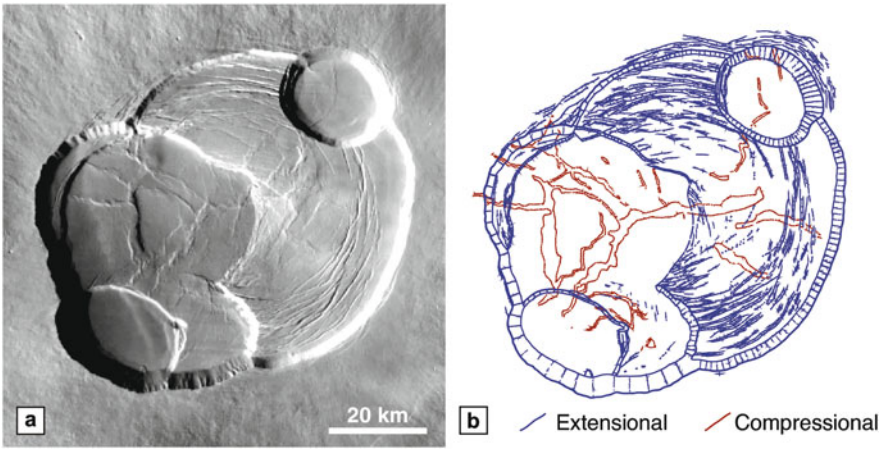

Fig. 8.7 Volcano-tectonic features at Martian volcanoes. (**a**) The summit of Olympus Mons (height: 21.2 km) is characterised by a nested caldera complex. (**b**) Structural map of extensional landforms (normal faults, grabens) on the caldera floor, including both extensional and contractional features. Source: (**a**) ESA/MEX/DLR/FU Berlin. (**b**) P. Kronberg and E. Hauber

Table 8.1 Volcanic eruption types (after Green and Short (1971), Wilson and Head (1994), Sigurdsson (2000))

Eruption name	Type	Description
Icelandic	Effusive	Fissure eruptions (gas-poor) producing great volumes of basaltic lava flows. Forms plateau basalts.
Hawaiian	Effusive	Eruptions from fissures, calderas, and pits; low-viscosity lavas that are erupted quietly, but can display some moderately explosive behaviour (lava or fire fountaining). Fire fountains erupt clods of lava that fall back still in molten state, form a lava pond, and continue as lava flow.
Strombolian	Explosive	Moderately explosive, rhythmic or continuous explosions; gas rises faster than magma in the conduit, bubbles form and coalesce until they disrupt the lava; ejection of lava clods that travel in ballistic trajectories and cool rapidly, falling down as bombs and unconsolidated scoria. Typical landforms are scoria cones.
Vulcanian	Explosive	Small to moderate-sized, short-lived (seconds to minutes) volcanic outbursts that eject material to heights <20 km. Discrete, violent explosions and ballistic ejection of blocks and bombs. More viscous lavas than in Strombolian eruptions can form a crust, under which gas pressure can build up until it is released violently.
Peléan	Explosive	Slow rise of high viscosity lava forms viscous dome or plug with solid carapace, catastrophic disruption, gas (and some lava) escapes from lateral flank openings, or destruction of plug; origin of a small convecting plume and pyroclastic flow (nuées ardentes) that can create block-and-ash flow deposits, made up from lava blocks with fine vesicles.
Plinian	Explosive	Magma disrupts into small fragments, efficient heat transfer to gas, entrainment of fine particles into rising gas stream, entrainment of ambient atmosphere; formation of very high convective ash plumes; creates (wind-transported) ash fall deposits, or, if the rising gas column becomes too dense and collapses, pyroclastic flow deposits (ignimbrites or ash-flow tuffs).

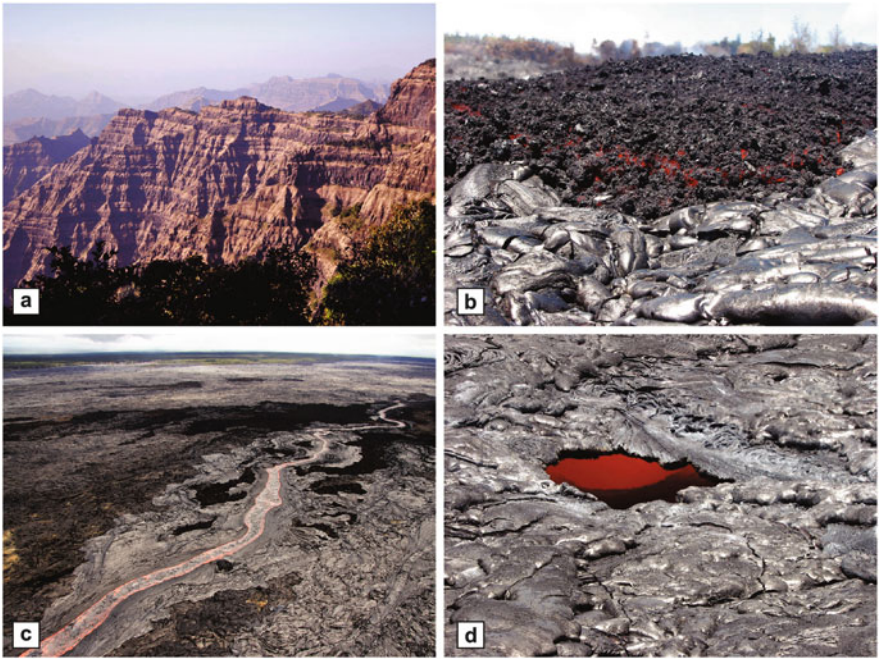

Fig. 8.8 Lava flow morphologies. (**a**) Trapp basalts in the Deccan province, India. Multiple flat-lying lava flows were eroded into mountains with a characteristic step-like topography. (**b**) An active *'a'ā* flow (background) advances on top of a recent *pāhoehoe* flow (foreground) in the March 2008 eruption of Kilauea, Hawaii. (**c**) Channelised lava flow on the northeastern flank of Pu'u 'Ō ō, Hawaii. (**d**) Skylight on a tube-fed lava flow on Pu'u 'Ō ō, Hawaii. Source: (**a**) Gerta Keller. (**b**) USGS. (**c**) Hawaiian Volcano Observatory, USGS. (**d**) USGS

e.g., the Deccan Traps in India or the Siberian Traps (the word trap comes from the Swedish term *trappa* for stairs and refers to the characteristic step-like topography of sub-horizontally layered flood basalts stacks (Fig. 8.8)). Several properties of lava flows have been used for their subdivision into separate types. The perhaps best known is the distinction between *pāhoehoe* flows and *'a'ā* flows, which may have identical chemical compositions yet display smooth and ragged surface textures, respectively (Fig. 8.8). It is commonly assumed that lava rheology, strain rate, and effusion rate govern whether a *pāhoehoe* or *'a'ā* flow is being emplaced.

Importantly, a single flow can transition from *pāhoehoe* type to *'a'ā* type, but not vice versa. Another morphologically relevant distinction is that between *channelised flows* and *tube-fed flows*. Open-channel flows typically display a central channel and lateral levées (Fig. 8.9), whereas tube-fed flows form an upper cooling crust, while the molten lava continues to flow underneath the crust. When this crust collapses, a so-called *skylight* (Figs. 8.8 and 8.9) is formed, and when the tube drains due to a ceasing lava supply, it turns into an empty lava conduit (often called lava tunnel). Tube-fed flows cool less efficiently than channelized flows due to the

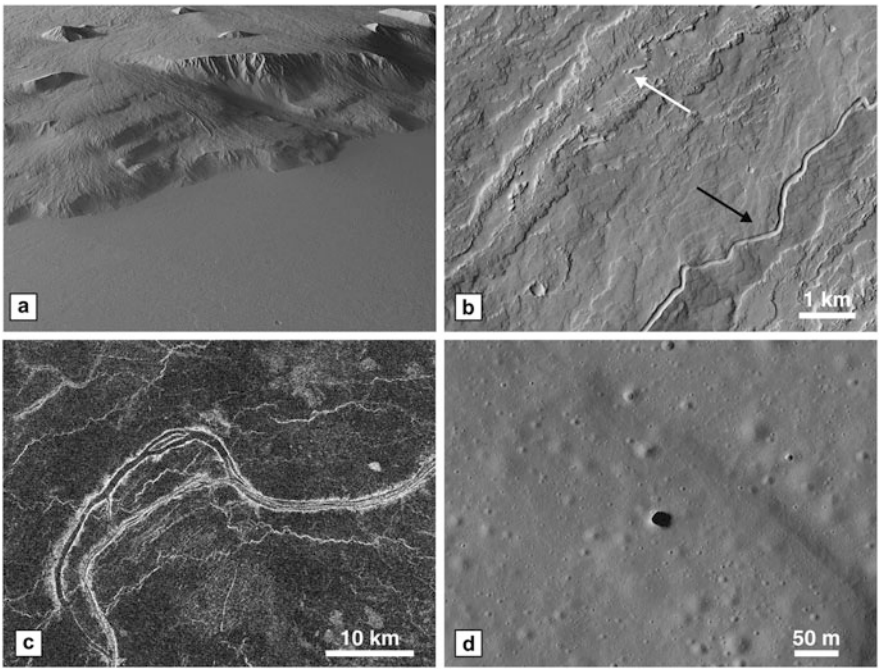

Fig. 8.9 Lava flow fields on the terrestrial planets. (**a**) Perspective view of the southeast flank of Olympus Mons. Successions of lava flows cascade down a 5-km high scarp. The lava fan at the base of the scarp is truncated by an extended smooth lava plain (foreground). Scene is about 70 km across. (**b**) The lava plain approx. 30 km NNE of Olympus Mons shows different lava flow types. *Lower right*: sinuous channel atop a lava tube, indicated by the *black arrow*; *upper left*: up to 2 km wide lava flow with characteristic levée and channel facies (*white arrow*). (**c**) Lava channel in Sedna Planitia, Venus. (**d**) Orbital view of a skylight on a lava-covered plain the Moon, see Fig. 8.8. Source: (**a**) ESA/MEX/DLR/FU Berlin, HRSC orbit 11524. (**b**) NASA/MRO/CTX, modified after Platz et al. (2015). (**c**) NASA/NSSDC/Magellan. (**d**) NASA/LRO/LROC/GSFC/ASU

protective effect of the solid crust, which minimizes radiative and convective heat loss. The cooling rate has an effect on lava flow length, and a lava flow stops when the heat loss coupled with increases in the crystal content force the viscosity and yield strength to increase to a point where motion is no longer possible (*cooling-limited flows*). Alternatively, lava flows can come to a halt before the maximum (cooling-limited) extent is attained when the lava supply stops (*volume-limited flows*). Composition, insulation and morphology being constant, flow length will positively correlate with effusion rate.

If the pressure release associated with the magma ascent leads to volatile exsolution and the formation of gas bubbles, these bubbles can ultimately disrupt the magma, leading to an *explosive eruption*. Depending on the nucleation depth of gas bubbles in the conduit, the viscosity, the ascent rate, and other factors, explosive eruptions can be more or less violent. Accordingly, several types of explosive eruptions are distinguished, ranging from moderately explosive Hawaiian eruptions

to highly explosive and potentially hazardous Plinian eruptions. Each eruption type produces certain landforms, although the distinction in remote sensing data is not always straightforward.

The interaction of magma with water and/or ice leads to rapid heating and water vapour formation. If this happens in a contained environment (e.g., groundwater heating in the subsurface), the vapour pressure can increase sufficiently to disrupt the containing material, leading to extremely violent *phreatomagmatic eruptions*.

Not all eruptions in aqueous environments are violent, however. Basaltic eruptions at mid-ocean ridges are typically quiet, because the high ambient hydrostatic pressure prevents violent magma degassing. A typical product of such non-violent subaqueous eruptions are pillow lavas, which may represent the most abundant volcanic rock type on Earth and are characterized by rounded (pillow-like) shapes and a rapidly quenched glassy rind. Other typical volcanic landforms on the ocean floors are individual volcanoes called seamounts. Subglacial eruptions represent a special type of magma-ice/water interaction. When the volcanic heat melts the overlying ice, a subglacial lake is formed and the eruption becomes subaqueous and may form pillow lavas first (if the lake is deep enough) and more explosive eruption products in shallow water later.

It is important to note that the style of an eruption can (and typically does) change during a single eruptive cycle. For example, an eruption can evolve from explosive to effusive or vice versa, and eruptions can change from an initial *wet* phreatomagmatic stage to a subsequent *dry* (e.g., Strombolian) explosive style. Accordingly, landforms created during an eruption may reflect these varying eruptive conditions, and any interpretation based on remote sensing data needs to acknowledge that the resulting landform assemblage may have a polyphase eruption history, and needs to consider the possibility of mixed deposits.

8.3 Tectonism: Driving Forces

Forces generate stress that generates deformation. In planetary crusts or lithospheres, contact forces, which act on rock surfaces, and body forces, which affect the entire volume of rock, coexist. Contact forces can be generated by a number of ways; for instance, on Earth, the weight of the cooling oceanic crust makes it enter into subduction zones, a process which generates traction forces (slab pull). Those forces generate the global system of stresses that drive the lithospheric plates and produce orogenic belts, rifts, and large strike-slip faults. On the other hand, body forces that can result in rock deformation are generated by gravity. Landsliding is an example of gravity tectonics.

Typically, the tectonic style is thought to be identical to the crustal deformation style. However, this style may not depend on the crust only. What is instead relevant is the portion of the planetary radius located above the thermal boundary layer above the convective mantle, if any. On Earth, this zone includes the uppermost part of the mantle and the crust, of distinct composition, which define the mechanical

lithosphere. Alternatively, and more simply, the lithosphere may also be defined as the region in which heat is transported to the surface mainly by conduction. Both lithosphere definitions give similar thickness estimates, though calculated differently. Another useful definition of lithosphere is based on the speed of seismic wave propagation. Whatever the definition on other planets, the absence of deep geophysical data is a major problem.

The tectonic style in the lithosphere varies with time, following the evolution of the heat flow, as high heat flow promotes ductility. It is, for example, not clear when plate tectonics started on the early Earth. On the other hand, Venus may have experienced plate tectonics in the past for which we do not see evidence anymore. In addition, a lithosphere that does not deform instantaneously may deform on geologic timescales by viscous relaxation. The style of deformation will then depend as well on the lithospheric composition, the water content, and the loading conditions. On Earth, where the lithosphere is made of two compositionally different components, crust and mantle, the question of how continental lithosphere behaves on the long term in response to these factors is not yet clear, in spite of the wealth of data available. Understanding the style and origin of tectonic activity of other planets is therefore a formidable challenge, in the absence of data regarding the composition and mechanical properties of the lithosphere, and taking into account the huge uncertainties on how the heat flow and water content of the crust have evolved through time.

8.3.1 The Tectonic Style of the Earth

Plate tectonics is the best documented example of global tectonics (Fig. 8.10). It currently operates on Earth only and requires several conditions. One is the presence of volatiles in the crust, which facilitates plate overthrusting in subduction zones, and for which evidence is testified since the Archean. Another one is that the plates are rigid and can deform at their margins only. The driving force is slab pull at subduction zones, but at smaller scale, body forces provided by density contrasts, for instance due to magmatic underplating, can influence intraplate stress conditions and hence the tectonic style. Although the view that plates are generally decoupled from the underlying mantle is frequently shared, it is increasingly considered that mantle flow contributes significantly to plate motion beneath oceans and kinematics at convergent plate margins.

The demonstration that plate tectonics operates on Earth was enabled combining information from two critical types of measurements. The symmetric distribution of magnetic anomalies in rocks of increasing age away from the mid-ocean ridges requires seafloor spreading on geologic timescales. GPS (Global Positioning System) and interferometric methods show that instantaneous plate displacement rates are equal to the displacement rates measured on geologic timescale, and demonstrate that internal plate deformation is negligible and that strain concentrates at sharp to diffuse boundaries. It is important to note that both methods require

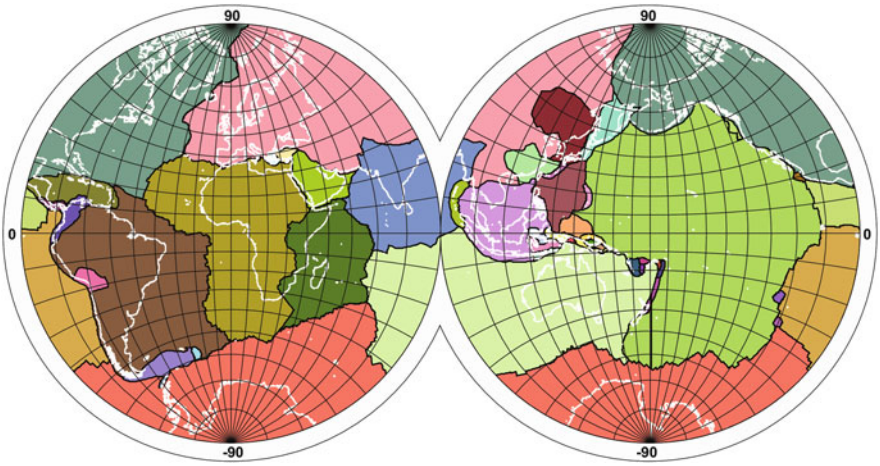

Fig. 8.10 Current tectonic plate boundaries on Earth, indicated with *dark outlines*. Continent outlines are indicated in *white*. Large plates occupy cratonic areas and oceanic basins, while (mostly) convergent zones host locally smaller plates, with very complex geometrical relationships between them. Source: Bird (2007) and later integration by Ahlenius, Nordpil, Bird, on GitHub

measuring absolute time differences at many measurement sites; without such information, plate tectonics could not be demonstrated. Such methods are currently not available on other planets; independent of the existence or absence of clues coming from other types of observations and methods. Determining whether plate tectonics ever operated on other planets is beyond current technological feasibility.

Before the onset of plate tectonics, a global tectonic system called *sagduction* tectonics or drip tectonics is thought to have prevailed, which produced the granite-greenstone belts that characterize Archean tectonics. *Granite-greenstone belt* is a descriptive term that accounts for the green colour of metamorphosed mafic to ultramafic rocks associated with granites and gneiss; attempts to understand their tectonic context has led to the definition of *ultra-hot orogens*, as opposed to the *collisional orogens* that are a product of plate tectonics. Structural evidence shows that crustal deformation of ultra-hot orogens resulted from sagduction, a vertical tectonics mechanism in which denser rocks at the surface sink into lighter rocks. Sagduction may be driven by an inverse density gradient between supracrustal rocks, mainly mafic or ultramafic, and underlying lighter, granite-gneiss rocks; or by lateral shortening and downward extrusion of the supracrustal rocks. This mechanism implies that the radiogenic heat production in the crust was high enough for the crust to be ductile over its whole, or most of its thickness.

8.3.2 The Tectonic Style of One-Plate Planets

Each terrestrial planet has its own tectonic style, resulting from differences in water abundance, internal heat (hence planetary mass), crustal thickness and temperature, and also their accretion history. The number of plates in the terrestrial sense is determined by the number of blocks that are internally rigid and peripherally deformed. Although this looks simple, this criterion asks fundamental questions such as the minimum dimensions of rigid portions of the crust that can be reasonably treated as plates, or rather *micro-plates*, and how much their kinematics needs to be different from the neighboring plates. Even today, determining the number of plates still critically depends on the availability of data. With the increasing number of kinematic observations worldwide, the number of lithospheric plates on Earth has increased from 6 in the first model of plate motions proposed by Le Pichon in 1968 to up to 50 (including many micro-plates) in 2014 (Fig. 8.10).

Perhaps because their geologic evolution has been simpler than that of the Earth, or perhaps because of the difficulties in identifying plates on other planetary bodies, they are usually all considered to be planets with only one plate. Should this view change in the future, and some other planetary bodies would possess a collection of plates, this would not necessarily be evidence of plate tectonics in the terrestrial sense, with accretion, subduction, collision, and strike-slip boundaries. We shall see that the Moon, Mercury, Venus, and Mars can be considered as one-plate planets (Fig. 8.11) but still each having their own style, dominated—as far as it is understood today—by impact cratering (Moon, Fig. 8.12), tidal despinning and secular cooling (Mercury), ultra-hot orogeny (Venus), and mantle plumes (Mars, Fig. 8.13).

8.3.2.1 The Moon

Tectonism on the Moon was (and is) dominated by impact cratering, directly or indirectly (Fig. 8.12). Impact cratering itself (Chap. 7) includes tectonic processes, the style of which depends on the rheology of the target. Impacts that are energetic enough to generate melting have two additional tectonic effects: related magma emplacement before cooling, and then again the creation of stresses once the magma has cooled and loaded the crust.

Infilling of the lunar craters and basins by basalt has various effects. Lava flooding is made possible by magma sheets (dikes and sills), the geometry of which may or may not be affected by the fractures produced by the impact and the gravitational modification of the crater or basin. Dikes and sills are hydrofractures that may produce significant deformation of the surrounding rocks if emplacement is shallow enough. Some narrow grabens that cut the surface of the maria have been interpreted to be a manifestation of such late-stage dikes.

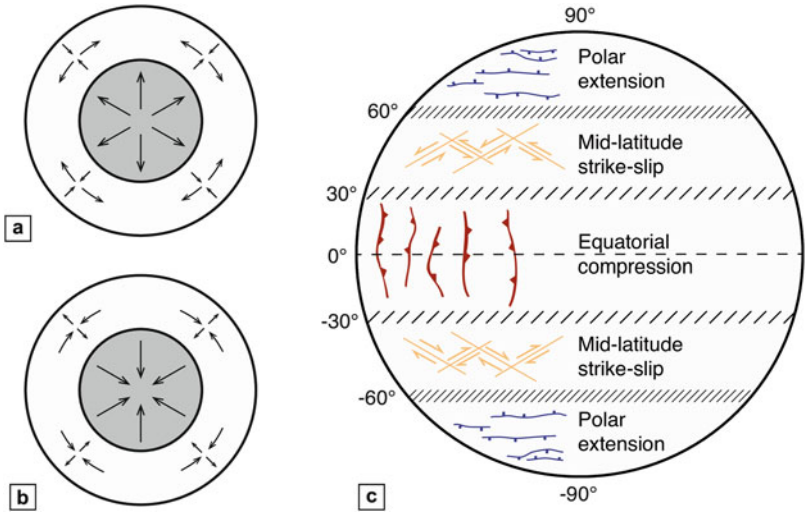

Fig. 8.11 Basic conceptual models of tectonic patterns of one-plate planets. Global radial and concentric stress patterns on one-plate planets due to secular heating and cooling: (**a**) Initial global heating and expansion, with dominant lateral tensional stresses in the crust. (**b**) Late-stage global cooling and contraction. (**c**) Predicted global tectonic pattern of a despinning planet. As the rotation rate decreases, the rotational flattening of the planet decreases and the polar and equatorial radii will increase and decrease, respectively. Correspondingly, different patterns of stress will develop at different latitudes. Source: (**a**, **b**) Redrawn after Solomon (1978). (**c**) Redrawn after Melosh (1977)

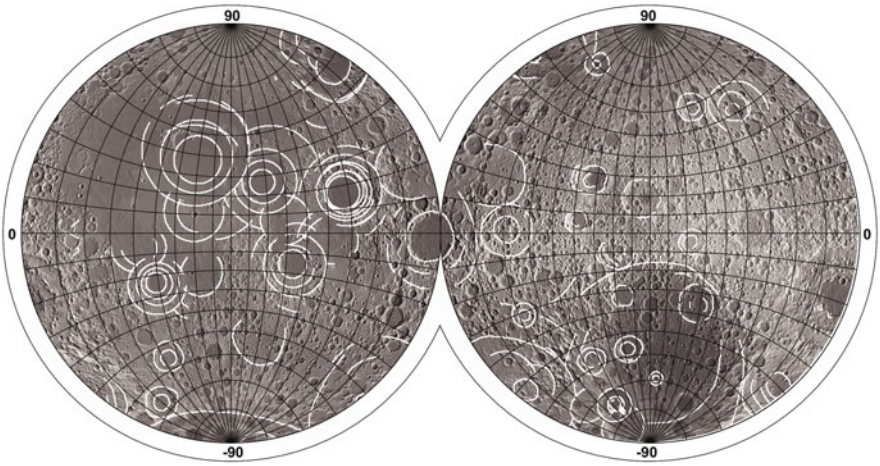

Fig. 8.12 Extensional structures related to large impact basins on the Moon. Contractional structures, mainly within basaltic maria, are not indicated on the figure. Source: redrawn after Geiss and Rossi (2013) based on USGS data from Lucchitta (1978), Scott and McCauley (1977), Stuart-Alexander (1978), Wilhelms (1979), Wilhelms and El-Baz (1977), Wilhelms and McCauley (1971)

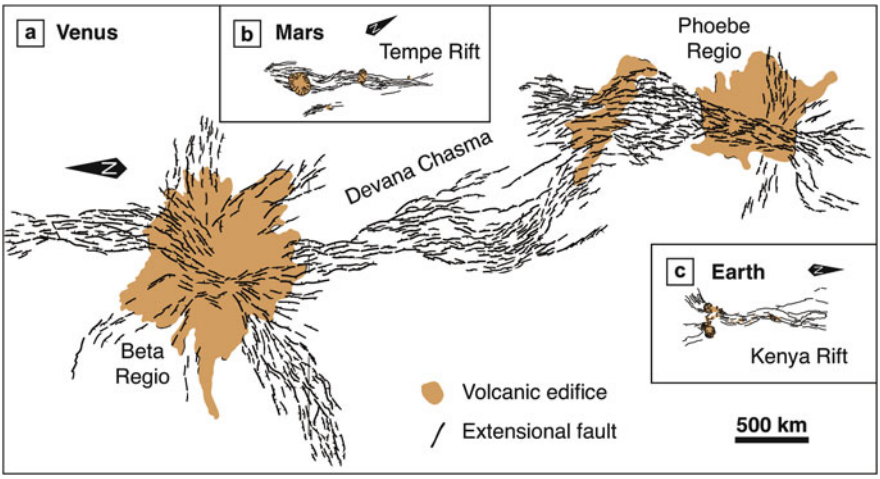

Fig. 8.13 Tectonic sketch map of Devana Chasma, a major rift-like extensional structure on Venus. (**a**) Devana Chasma is comparable in size and structural architecture to (**b**) other extensional systems on Mars as well as (**c**) terrestrial continental rifts such as those in the East African Rift System (e.g., Kenya Rift). Note the large volcanic centres of Beta and Phoebe Regio, which are linked by the rift system. Source: Mapping by P. Kronberg, plus, modified from Hauber and Kronberg (2005)

Lava pool cooling in impact craters and maria has a contractional effect on the surface at the center of the maria, generating wrinkle ridges. The contraction is compensated by extension at grabens that accommodate thermal stress release at the boundary between the cooling lavas inside the maria and the surrounding cool crust.

8.3.2.2 Mercury

One-plate planets may display homogeneous or heterogeneous crustal deformation; Mercury is the one displaying the highest degree of homogeneity. Due to its proximity to the Sun, the crust, which has an estimated average thickness of 35 km $\pm$ 50%, is thought to be dry, preventing any large-scale tectonic motions.

Although the surface of Mercury looks approximately like the surface of the Moon, it differs by the similar, mainly basaltic, composition of the impact structure fill and floors and their host crust exposed around them, as opposed to the difference between lunar basaltic mare and anorthositic highlands, and by the widespread surface deformation by contractional structures. Extensional structures are very rare, though not absent. The abundant contractional features (wrinkle ridges and lobate scarps) do not display an obvious global tectonic fabric, although it was recently shown that compression at mid-latitudes is preferentially E-W oriented. Contraction is sometimes associated with strike-slip deformation. These structures are significantly overprinted only by the Caloris impact basin. There has been

little variation in the possible sources of deformation that were investigated since the Mariner 10 flybys (1974–1975), including two global types of events, secular planetary cooling and early tidal despinning (Fig. 8.11).

The effect of secular cooling on the observed global contraction of Mercury has been little questioned, as cooling produces a net decrease of planetary size, necessarily resulting in net surface contraction while tolerating strike-slip faulting. Tidal despinning has been questioned because in the absence of any stress field of other origin active at the same time, the predicted associated deformation also includes extensional faulting. However, when secular contraction and tidal despinning are simultaneously active, the predicted tectonics is fully compressional, with the mid-latitudinal preferential E-W contraction favored, and contraction of random orientation close to the pole is predicted, as observed.

The Caloris region displays additional distinctive tectonic patterns, especially the remarkable graben system of Pantheon Fossae that radiates from the Caloris Basin center and probably formed by surface uplift, for instance following the growth of an underlying intrusion. Based on structural similarity between Pantheon Fossae and structures such as novae on Venus, it was suggested that they formed coeval with dike emplacement at depth.

8.3.2.3 Venus

Despite its similar size and proximity to Earth, which suggests that its bulk composition, inherited from accretion, is also similar at first order, Venus currently lacks plate tectonics and is thought to be a one-plate planet. The high temperature of the crust of Venus, which, in contrast to Mercury, is not due to the proximity to the Sun, but to the high temperature of the atmosphere, does not allow water to be stable within the crust either. This is one of the reasons to explain the absence of plate tectonics. Why there is no water on present-day Venus is not fully understood. Furthermore, the hot surface temperature of ~460 °C was recently shown to cause potential crustal discontinuities to heal, preventing any plate boundary from even initiating.

Volcanic plains occupy more than half of the Venusian landscape. They are interrupted by broad volcanic rises, huge (*crustal*) plateaus, and the Ishtar Terra region. The volcanic plains, and some volcanic rises, include chasmata and coronae structures with diameters ranging from 60 to several hundreds of kilometers (with the exception of Artemis Corona: 2600 km), and a stellate variant called novae. Venusian coronae have no recognized equivalent in the Solar System. Additionally, the volcanic plains display various types of small-scale deformation features (wrinkle ridges, fracture fields, polygons). The volcanic rises include rift zones (Fig. 8.13), and the crustal plateaus display tesserae, a complex superimposition of folds, faults, and lava flows reminiscent of a parquet geometry. Tesserae are, however, not always associated with plateau topography. Ishtar Terra is a huge, unique high area that includes tesserae in the east, the Maxwell Montes orogen in the center, and in the west a high, smooth plateau, Lakshmi Planum, surrounded by compressional belts.

Not all the tectonic features may be due to internal dynamics; the polygonal and other deformation features of the volcanic plains have been shown to be possibly a consequence of climate warming, from a period of moderate greenhouse to the present atmospheric state. Given the difficulty of rigid plate motions on present Venus, the tectonic features listed above are interpreted in terms of dominantly vertical motions, i.e., variants of upwellings and downwellings.

In contrast, coronae are interpreted as ovoid features forming in the crust in response to buoyant diapirs in the mantle. Coronae have a variety of shapes, from domes to collapsed plateaus, and are surrounded by trenches, which can be reproduced by models of diapiric rising and subsequent spreading in response to cooling. Novae are radiating grabens that are not associated with ovoid structures and trenches and correspond to early stages of corona formation. They are associated with pit craters, and other morphologies of likely magmatic origin (sometimes lava flows), indicating that they are the surface expression of underlying dike swarms. As a consequence, many coronae (47%) are also associated with radial grabens, and their mapping can be used to reveal local stress fields.

Five main rifts zones have been identified on Venus, some of which form a large equatorial rift zone system. Some of the Venusian rifts are structurally very similar to terrestrial continental rifts associated with flood basalt provinces such as the East African Rift System. Analysis of rift topography at Devana Chasma (Fig. 8.13), one of the main equatorial rifts, similarly revealed rift flank uplift.

8.3.2.4 Mars

It has been a long debate how many tectonic plates constitute the Martian lithosphere, with terrestrial-type plate tectonics interpretations proposed since the Mariner 9 mission. All these interpretations fall short at global scale, because each defined plate boundary has planet-wide implications on planets of constant radius. For instance, removing a portion of the lithosphere in a subduction zone needs compensation by forming new lithosphere elsewhere, in a spherical puzzle in which each plate boundary kinematics has influence on the kinematics at the other boundaries. One of the most interesting, and perhaps the most fundamental, issue in Martian tectonics is the nature of the so-called *hemispheric dichotomy boundary*, a topographic transition zone that separates the northern plains and the southern highlands. The topographic difference between the lowlands and highlands is ~2 km, and the transition is marked by a mean slope that does not exceed 3–4° over a distance of a few hundred km to 3000 km. The age of the dichotomy is not very well constrained, although it is considered likely that it formed already very early during the pre-Noachian, making it the oldest large-scale morphologic feature of Mars (Chap. 11). There is no evidence of faulting that could have been involved in the formation of the dichotomy. Gravity and topography data inversion has been used to propose crustal cross sections in which the topographic decrease toward north is associated with a rise of the crust-mantle boundary, which suggests that in a descriptive sense, Mars may be considered a two-plate planet. In a kinematic sense, however, it is highly unlikely that the lowlands and highlands once moved

with respect to each other along the boundary. Because of the irregular shape of the dichotomy boundary, such movements would have deformed it intensely, which is not observed. Kinematically, Mars is today best considered as a one-plate planet.

The Tharsis volcanic-tectonic bulge (Fig. 8.14) was built on this dichotomy boundary, which it broadly masks. The bulk volume of Tharsis, the largest known

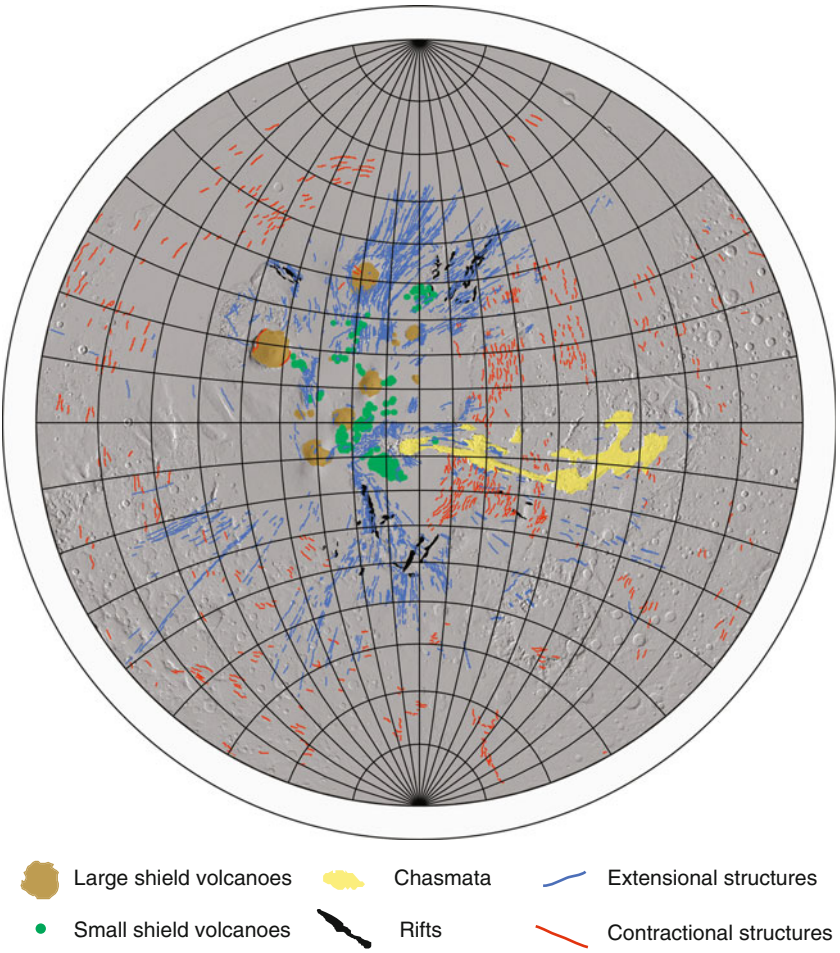

Fig. 8.14 Tectonic sketch map of Tharsis, the largest volcanic province in the Solar System. The huge topographic bulge dominates the western equatorial hemisphere of Mars. It is characterized by very large shield volcanoes (*brownish colors*) and hundreds of smaller volcanic vents (low-shield clusters are shown in *green*), several sets of long and narrow grabens (*thin blue lines*) that radiate outwards from several centers, and a concentric set of wrinkle ridges (*thin red lines*). Volcanic loading of the lithosphere is likely responsible for the concentric tensional stress and the radial compressive stress. A few large and complex extensional features (in *black*) are comparable to terrestrial continental rifts. The ~3000 km-long trough system of Valles Marineris (*yellow*) is controlled by Tharsis-radial trends and was probably formed by a combination of extension, collapse and erosion. Source: Modified from Hauber and Kronberg (2001)

volcano-tectonic province in the Solar System with a diameter of about 5000 km, formed early in the history of Mars. Its exact formation age is somewhat debated, but its activity, at least locally, continued until relatively recent times (Chap. 11).

Many extensional structures (faults, usually forming grabens, and tension fractures) are radiating from the Tharsis central area, and oriented perpendicular to contractional structures such as wrinkle ridges (Fig. 8.14). These extensional structures may be the surface expression of giant radiating dike swarms, analogous to dike swarms in ancient cratons on Earth.

The intensely deformed zones that predate the Tharsis volcanic flows are organized in *rift zones* that were later reactivated in the context of Tharsis evolution. Rift zones are extensional features that are geometrically and kinematically (but not geodynamically) very similar to terrestrial rifts.

Valles Marineris is a collection of elongated troughs, some interconnected, that extends over 3000 km along strike and encompasses a region 800 km wide. Elongation is parallel to the surrounding Tharsis radial grabens, which are sometimes cut by the trough walls, showing that tectonic extension participated in their formation.

The southern troughs are linear, and triangular-faceted spurs suggest a tectonic origin. The northern troughs, in contrast, are partly or completely closed and have distinctly non-linear plan-view outlines, suggesting dominating (and not fully understood) collapse and erosional processes for their origin. The relative role of tectonic extension vs. collapse in Valles Marineris and related chasmata remains one of the most puzzling enigmas of the geologic evolution of Mars.

8.3.2.5 Icy Satellites and Kuiper Belt Objects

The surfaces of some icy satellites in the outer Solar System display many tectonic features that are not found on the rocky terrestrial planets. Their icy shells are deformed by tidal stresses induced by their parent planet and by other satellites, and their surfaces bear clear records of these stresses (*cryotectonics*). The nature and magnitude of deformation depends on orbital parameters (e.g., eccentricity) and the satellites' obliquity (i.e., the tilt of their rotational axes with respect to their orbital planes) as well as on the thickness of the icy crust, and the rheology of the underlying material. In some cases (e.g., Europa and Enceladus) it seems likely that global subsurface oceans underlie the rigid outer shells, and fractures penetrating the entire crust would perhaps enable the ascent of water to the surface (*cryovolcanism*). As tidal stresses may be considered to be due to external forces, they will not further be discussed in this chapter on endogenic processes.

When the New Horizons spacecraft delivered the first detailed images of the surfaces of Pluto and Charon, one of the biggest surprises was the apparently young age of some surfaces. Obviously, tectonic and perhaps cryovolcanic processes were much more active in the recent geologic history of these Kuiper Belt objects than previously thought. At the time of writing, the spacecraft still transmits data of its flyby, and only preliminary results are available. One of the early findings is that convection in a layer of nitrogen ice forms characteristic polygons with a diameter of 10–40 km. The lack of impact craters in this region, informally called Sputnik

Planum, implies continuous and relatively quick resurfacing. These results add to the emerging view that some small and icy bodies in the Outer Solar System are geologically dynamic bodies (Chap. 13).

8.4 Magmatism and Volcanism: Driving Forces

8.4.1 Igneous Volcanism

On Earth, three main geodynamic settings associated with volcanism may be generally distinguished: volcanism along convergent and divergent plate boundaries, and intra-plate volcanism. Magmatic activity concentrates along plate boundaries and so do volcanoes. The circum-Pacific *Ring of Fire* is the most active and widely known example today. It is mainly located at convergent plate boundaries and associated subduction zones. Even more important in terms of flux and volume, new crust is continuously formed at plate boundaries along the mid-oceanic ridges in a process called sea floor-spreading.

On divergent plate margins, lithospheric plates move apart. As a consequence, the overburden load of the thinned crust above the mantle decreases, hence mantle material can move upward to fill the space between the moving plates. The adiabatic ascent and the decrease in lithospheric pressure enable the generation of adiabatic melts without the need of additional heat—a process known as decompression melting of mantle material. This partial melting generates basaltic magma with low silica contents. An important consequence of volcanism at or along ocean ridges is *hydrothermal activity* where cold seawater is heated and enriched in various chemical constituents.

At convergent plate margins, an oceanic plate is subducted underneath either a continental or oceanic plate, and the subducted oceanic crust descends into the mantle. The delivery of wet crust, carrying water-rich sediments with it, into the hot mantle causes partial melting of the mantle above the subducted plate, as water reduces the mantle liquidus temperature. Melt generation, ascent, and subsequent fractionation and differentiation towards the surface commonly produce high-silica magmas with substantial volatile contents (mainly H_2O, CO_2). Such magmas would often erupt explosively.

Based on our current understanding, all other terrestrial planets are one-plate planets characterized by so-called stagnant lid regimes (Chap. 10). As a result, volcanism, i.e. magma generation and extrusion is typically explained by heat transfer through *mantle plumes* or large-scale mantle upwellings.

Mantle plumes are thought to be hot, buoyant diapiric upwellings of solid mantle material that rise through a planet's mantle in a convective mushroom-like manner, carrying heat from the depth, most often assumed to be the core-mantle boundary, to the surface over a long period of time. Once the top of the plume is close to the surface, the drop in pressure may allow the temperature to cross the solidus causing the generation of a large volume of melt in the upper mantle. These melts may

then intrude into the crust or erupt onto the surface to form flood basalt lavas. The locations of such volcanic activity are called *hot spots*. The melting associated with mantle plumes generally produces large volumes of melts periods of time; however, there is a decrease in magma volumes over time as the plume may laterally spread out in the mantle beneath the crust.

On Venus and Mars, magmatism may have been largely associated with mantle plumes, and the style of volcanic activity may reflect this *plume tectonic regime* and would be analogous to hot spot volcanism on Earth, enabling the formation of large-scale volcanic regions with giant volcanic edifices, such as in the Tharsis and Elysium provinces on Mars.

Under certain circumstances, volcanism may dominate the heat transport on planetary bodies from their interior to the surface (Chap. 10). If this is the case, a body enters a *heat-pipe mode* where heat is transferred to the surface by magma ascending throughout the lithosphere in narrow channels. This may be the case for Jupiter's moon Io, the most volcanically active body in the Solar System. The mechanism that produces the internal heat of Io is widely agreed to be dissipation of tidal energy from interaction with Jupiter, induced by the orbital resonance 4:2:1 between Io, Europa, and Ganymede.

The mechanism of cooling by heat-pipes was also proposed for early Earth before plate tectonics developed, as well as for other terrestrial planets early in their evolution.

8.4.2 Non-igneous Volcanism

Volcanism, however, is not restricted to the production of silicate magmas (igneous volcanism). Several icy moons in the Outer Solar System (mainly moons of gas giants and dwarf planets) display evidence for heat transport by cold fluids and volatiles. Such phenomena are referred to as *cryovolcanism*, because they occur in environments with extremely low temperatures. During cryovolcanism, explosive eruptions or effusions are triggered by fluid water and/or aqueous solutions of several other chemical components such as ammonia, methane, nitrogen, hydrocarbons etc. Active cryovolcanism was observed by spacecrafts on Saturn's mid-sized icy satellite Enceladus and on Neptune's largest moon Triton (Chap. 12).

Another type of non-igneous volcanism is *mud volcanism*. During this process a mixture of liquid water and sediment is mobilized and extruded at the surface where it can form a *mud volcano*, which may only be a few metres high. This mobilization of soft sediment can be driven by gases (e.g., CH_4, CO_2 and others); however, mud volcanoes may also be formed by non-volcanic processes, especially in sedimentary basins. For example, kilometre-scale mud volcanoes in Azerbaijan are associated with hydrocarbon deposits from which methane is released into overlying sediments. Additionally, mud volcanoes may be formed also by tectonic processes such as overpressurization in compressional settings. Mud volcanism is a common phenomenon on Earth, and was also suggested to explain the origin of several types of small pitted cones on Mars.

8.5 Magmatic Activity

8.5.1 Composition

The primary product of mantle melting is magma with basaltic composition, and therefore the majority of volcanic rocks on Earth are basalts. Basalts are generated in all volcanic settings, but they are especially dominant at mid-oceanic ridges and intraplate hotspots (flood basalt provinces). More evolved magmas are typically associated with convergent plate margins, where new granitic crust is being produced. On other planets, the dominance of basaltic igneous rocks is even more pronounced. Chemical and mineralogical information, morphologic characteristics, and available samples all point to basaltic compositions for most volcanic rocks on Venus, Mars, Mercury, and the Moon. Nevertheless, there are exceptions. It is known that there are felsic rocks on the Moon, e.g., granites and rhyolites, although they are rare and volumetrically insignificant. It has been speculated that the tessera terrains on Venus may have evolved compositions and may be the equivalent to terrestrial granitic continents. Recently, observations by orbiters and rovers hinted at evolved compositions on Mars, ranging from quartz diorite and granodiorite to trachy-andesite and trachyte. These rocks would represent evidence for low-density crust on early Mars (Fig. 8.15). The discovery of the low-pressure, high-temperature

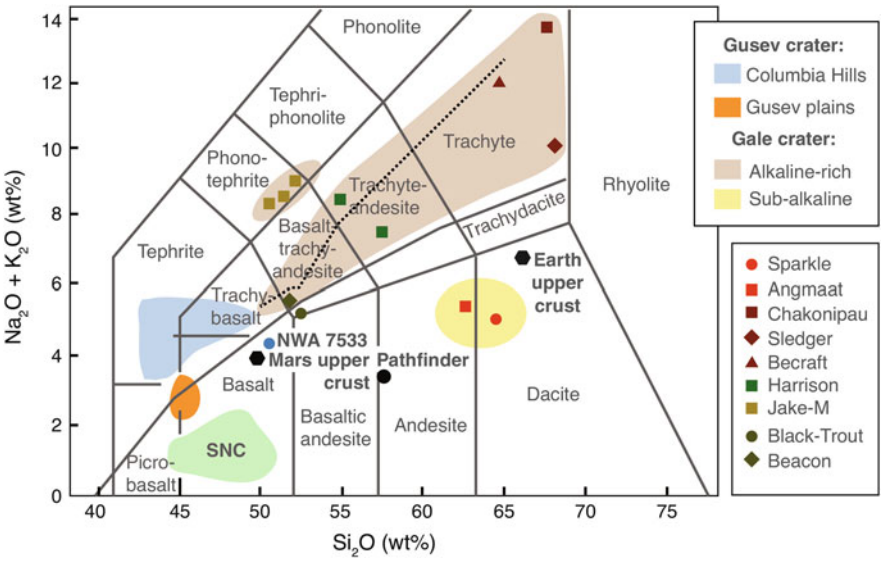

Fig. 8.15 TAS (Total Alkali versus Silica) diagram showing the range of observed crustal compositions of Mars. The *symbols* in the legend on the lower right correspond to samples analysed by the MSL rover, Curiosity, in Gale crater. Such trachy-andesitic, trachytic, and dacitic compositions may represent an early Martian crust that may be compared to continental-type compositions on Earth. The SNC green field includes all Martian meteorites except the Noachian breccia NW7533. Source: Modified from Sautter et al. (2015)

(~870 °C) SiO_2 mineral, tridymite, by the MSL Curiosity rover is the first in situ evidence for silicic volcanism and provides further evidence for evolved magmas on Mars.

More exotic compositions such as sulfur flows were thought to be possible on Io, but the observed extremely high eruption temperatures (~1,500 K) on Io basically rule out this possibility, as sulfur volcanism would not generate temperatures above roughly 700 K. It is now generally thought that lavas on Io consist mainly of superheated basalts (emplaced after rapid ascent from a deep, high-pressure source) or, alternatively, of cooled ultramafic komatiites. Nevertheless, the actual composition is unknown and remains one of the biggest critical questions in the study of Io.

8.5.2 Plutonism/Intrusions

Melts, once formed, are typically less dense than their parent materials (water is an exception), hence they have a tendency to rise towards the surface. The behaviour of magma (melt plus crystals) near the surface is then dependent on the contrast between the density of the magma and the surrounding rocks. If the magma is less dense than the surrounding rocks, it may reach the surface and erupt. However, if the magma's density is similar to that of the surrounding rocks, the level of neutral buoyancy is reached, causing the magma to stall where it typically forms magma reservoir(s). Over time such reservoirs may fully crystallise and form intrusive igneous bodies of various sizes (e.g., batholiths, laccoliths, etc.).

On Earth, the relative proportion of intrusive versus extrusive rocks (the I/E ratio) is typically estimated to range from 5 up to 40, although this depends on the tectonic environment and the properties of the crust through which magma is rising. Specifically, new crust produced at mid-ocean ridges has a higher than usual percentage of intrusive rocks. Nevertheless, it is also expected that the I/E ratio increases with increasing crustal thickness: the thicker crust acts as a filter, therefore a smaller fraction of magma reaches the surface. For instance, for Io or the Moon, which have very thick and very cold lithospheres, it is expected that the I/E ratio would be possibly much higher than on Earth. As an example, it is commonly believed that the concentration of mare basalts on the lunar nearside is a function of the different crustal thicknesses, with the crust being thicker on the lunar farside. Large magma intrusions have an important effect on the magnetization of the crust, because they may thermally demagnetize the shallow crust beneath volcanoes through heat conduction and associated hydrothermal activity.

Some structural features on Mercury's surface such as floor-fractured craters and radiating grabens may suggest intrusive activity, but overall the evidence is limited. This may be due to the dominance of contractional deformation, which may tend to prohibit the emplacement of shallow intrusions such as dikes and sills. On the Moon, intrusive bodies are thought to exist in association with grabens, and in floor-fractured craters, interpreted to be a surface manifestation of sills beneath them.

On both bodies, Mercury and the Moon, dike intrusions might originate from large depths, and the existence of shallow magma reservoirs does not seem to be common. On Venus, many surface features are likely associated with intrusions; hence, many large and deep magma reservoirs or diapiric upwellings together with both shallow and deep intrusions are expected to be present. On the icy satellites, non-silicate intrusive bodies were proposed as an explanation for linear ridges on Jupiter's rocky-icy moon, Europa.

8.5.3 Effusive Volcanism

Evidence of effusive volcanism is common on all terrestrial planets and Io. For example, the majority of volcanic terrains on the near side of the Moon is associated with extensive lava plains infilling lowlands and large impact craters. They are known as lunar *maria* and they resemble lunar equivalents of terrestrial *flood basalts*. The mare basalts represent by far the majority of volcanic rocks on the Moon, yet they occupy only about 17% of the lunar surface, and have a volume of 10^7 km^3, or 1% of the crustal volume. They were mainly formed early in lunar history, and the peak flux occurred during the Imbrian period about 3.85–3.2 Gyr ago. Nevertheless, younger mare flows have been identified, and recent observations suggest that endogenic activity on the Moon persisted until much later and may be as young as one billion years, or even less. On Mercury, 27% of the planet is covered by flat areas termed as *smooth plains* which have many similarities with lunar maria and are mostly located in the northernmost hemisphere and around some large impact basins such as Caloris, Beethoven, and Rembrandt. These plains do not display any unambiguous volcanic landforms, but the morphologic and structural characteristics are consistent with a volcanic origin. The analysis of crater populations that are either buried by or superposed on the smooth northern plains suggests that the volcanic plains have been emplaced over a geologically relatively short time interval of perhaps less than 100 Myr. They represent a minimum volume of 4×10^6 to 10^7 km^3 and are comparable to terrestrial Large Igneous Provinces (LIP). It seems likely that they formed in a single, voluminous volcanic event associated with extensive partial melting of Mercury's mantle at around the same time (between $\sim$3.7 and 2.5 Ga), similar to the major episode of mare volcanism on the Moon.

On Venus, about 70% of the surface is covered by volcanic lava plains, indicating that effusive volcanism played a major role for the formation of the currently visible surface. The Venusian surface has a relatively small amount of impact craters suggesting a young age, therefore, an intense period of effusive activity had to occur relatively recently, between $\sim$1 and 0.3 Ga. The more or less random distribution of impact craters on Venus also suggests that volcanic resurfacing was a truly global process. It is not known what kind of endogenic process might have been responsible for such an intense pulse of new crust formation. On Mars, flood basalts are also common; they occur as lava plains surrounding the volcanic centres of Tharsis and Elysium, but are globally widespread as so-called ridged plains. They seem to be

possibly related to a period of intense volcanic resurfacing in the Early Hesperian, about 3.6 Gyr ago.

Evidence for localized volcanic centres of various sizes is also common on Mars and Venus. For example, the entire surface of Venus is speckled with large central shield volcanoes, such as Maat Mons or Sif Mons, which are associated with lava flows emanating from their vents. The surface of Mars is characterized by large effusive volcanoes such as Olympus Mons (Fig. 8.7), but also by small effusive volcanic centres, e.g., large portions of Tharsis and Elysium are covered by young extensive lava flows (Fig. 8.9).

8.5.4 Explosive Volcanism

When ascending magma is fragmented into small volcanic particles, called *pyroclasts*, an *explosive eruption* occurs. Basically, there are two possibilities how explosive eruptions originate and how magma might be fragmented. One can be considered as a dry process, in which the eruption is driven solely by volcanic gases originally dissolved in the melt. It occurs when magma ascends relatively quickly from depth and is accompanied by rapid decompression. On Earth, the most abundant volcanic landforms produced by this type of moderately violent explosive activity are *scoria cones*, small conical volcanoes composed of unconsolidated scoria material. Evidence for such magma degassing is known from the Moon, Mercury, Io, and from Mars. Venus has a very dense atmosphere which hinders rapid degassing, and it is debated whether explosive volcanism occurs on Venus or not.

If volcanic gases are trapped in magma by some form of blockage of the conduit, e.g., by a viscous lava dome, pressure inside the volcano builds up and can eventually blast the blockage away. The sudden release of pressure causes rapid gas expansion, intense magma fragmentation, and hence creation of fine-grained volcanic ash. On Earth, such violent explosive eruptions are common for composite volcanoes rising above subduction zones. On Mars, some of the old highland volcanoes (e.g., Hadriacus, Tyrrhenus, and Apollinaris Montes) may have been formed by a similar process.

The second possibility how magma can be fragmented are *wet* (*phreatomagmatic*) eruptions which occur when magma of all types is mixed with an external source of volatiles, e.g., on Earth, seawater or groundwater. The basic principle of this interaction is rapid heat transport from magma to the volatile component, leading to volatile vapourisation, steam expansion and pressure build-up, and explosive magma fragmentation. On Earth, small volcanoes such as *tuff rings* and *tuff cones* are the most abundant volcanic landforms associated with this process. Which specific type of deposit is formed by a phreatomagmatic eruption depends, among other factors, on the amount of water available to drive the magma fragmentation (water-magma ratio). Phreatomagmatic eruptions should have occurred on Mars and Io, where the contact between lava and underlying frozen water or SO_2 may have led to explosive volatilization and to the formation of plumes.

8.5.5 *Environmental Effects*

The specific planetary environment plays a crucial role determining the type of volcanic activity and hence the shape of volcanic landforms. The environment consists of a complex array of factors, where each of them contributes differently on various aspects of volcanic activity. For example, the gravitational acceleration affects the depths at which magma reservoirs are likely to be formed, but it may also affect the geometry of magma chambers and feeder dikes. With decreasing gravitational acceleration, the sizes of these bodies need to increase to enable buoyant ascent through the country rock. Additionally, gravity affects the distance to which volcanic pyroclastic particles may be ejected in an explosive eruption, thus the ejected particles travel farther from the vent with decreasing gravitational acceleration (Fig. 8.16).

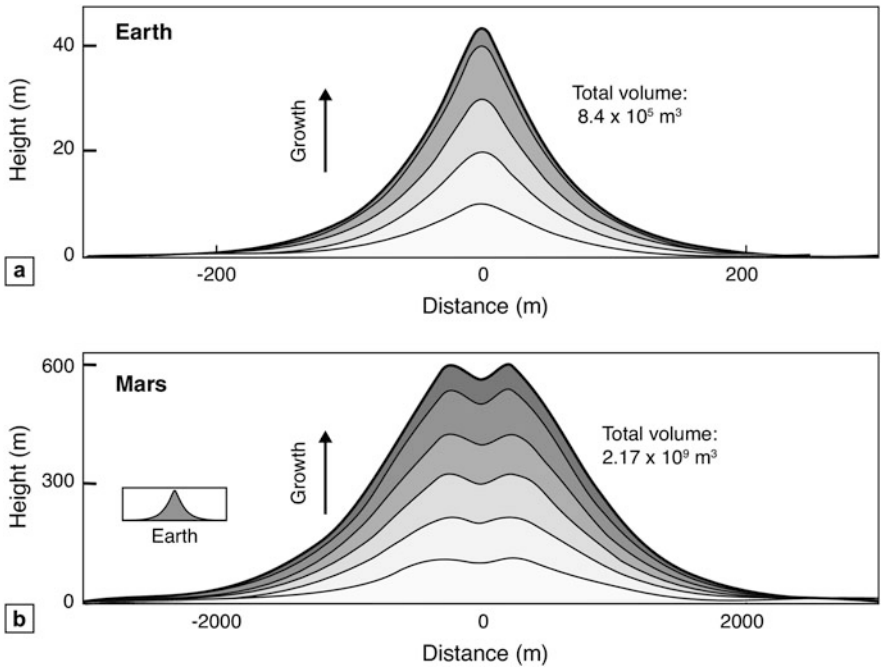

Fig. 8.16 Example of effects of environmental conditions on volcanic eruptions. (**a**) Evolution of a terrestrial scoria cones until the angle of repose (30°) is reached. *Increasing darkness of the fill* indicates the gradual growth of the cone, the *thick solid line* shows profile when angle of repose is reached, corresponding to a height of ~40 m and a volume of about 8.4×10^5 m^3. (**b**) The equivalent height of a cone on Mars is ~600 m, and the angle of repose is reached when the volume of deposited material is about 2.17×10^9 m^3, four orders of magnitude larger than on Earth. Note the dramatically different cone sizes on Earth and Mars when the angle of repose is first reached, a consequence of the much lower gravitational acceleration and atmospheric density on Mars, which enables ejected particles travelling much farther from the vent and thus covering a much larger area. Source: modified from Brož et al. (2014)

Another factor is the atmospheric surface pressure, or the absence thereof. On Earth, subaerial eruptions occur at an ambient pressure of ∼0.1 MPa, but this value is quite different on other terrestrial bodies. For example, on Venus, the pressure may be up to ∼9.2 MPa (92 bar), preventing volatiles to be easily released from magma. At the other extreme end of ambient pressure conditions, near vacuum exists on Mercury, the Moon, and Io. The absence of an atmosphere capable of inhibiting the exsolution of volcanic gases means that there is a significant drop between the pressures in the ascending magma column and the surrounding environment. Therefore, volcanic gases expand violently and rapidly when magma reaches the surface, causing ejection of magma clasts with high speeds. The ejection of volcanic particles is then characterized by the formation of an umbrella-shaped plume, as such environments do not support the rise of buoyant convective ash plumes as it is common on Earth. Actually, active extraterrestrial volcanism was first detected when a ballistic plume was seen against the dark space in Voyager images of Io (Fig. 8.1).

8.5.6 Outgassing

Outgassing is an especially important phenomenon shortly after the planetary accretion and core formation, when the surfaces of terrestrial planets may be covered by a global magma ocean (Chap. 11), from which large quantities of gases may have been released. This first period of intense outgassing produces a primary atmosphere, which over time is eroded and removed by solar wind and cosmic radiation. When in a later stage of planetary evolution magma reaches the surface, large amounts of dissolved gases, mainly water vapour, carbon dioxide, and sulphur dioxide can be released into the environment, a process called *outgassing*. Together with the impact delivery and the reactivation of subsurface reservoirs, volcanic outgassing is the main source of volatiles for the formation of a secondary atmosphere. Volcanism has a profound impact on the evolution of planetary atmospheres, and potentially influenced the early climate on Earth, Venus, and Mars. On shorter timescales, volcanism is historically known to have significant effects on the climate. On Titan, the atmosphere probably originated from ammonia which was released on the surface by massive volcanic activity, and was later converted to nitrogen by ultraviolet photolysis.

8.6 Volcanic Characterization of Solar System Bodies

8.6.1 The Moon

Compared to Venus and Mars, the inventory of volcanic landforms on Mercury and the Moon appears relatively limited, both in terms of diversity and abundance. On

both bodies, extensive plains of basaltic lavas are the dominant volcanic features. On the Moon, these plains (*maria*) consist of mare basalts, which are well visible from the Earth as large dark areas. Mare basalts form very leveled (equipotential) surfaces and lavas could flow over very long distances (~1200 km), which is thought to be a result of their very low viscosities. Vent structures are rarely observed, and typically the lava flows cannot be traced back to their sources. A particular type of landform associated with mare volcanism are *sinuous rilles*, meandering channels that often start at a circular or elongate depression and gradually disappear into the lava plains of the maria. The lengths can range from a few kilometers to more than 300 km. There is a consensus among many researchers that the channels are formed by a combination of mechanical and thermal erosion of the underlying substrate.

Regional dark mantle deposits, mostly associated with uplands adjacent to younger mare regions, are interpreted to be pyroclastic materials, probably dispersed over tens to hundreds of kilometers by Hawaiian-style fire fountaining driven by continuous gas exsolution. Pyroclastic glass formed by lava fountaining of gas-rich, low viscosity and Ti-Fe-rich basaltic magmas is volumetrically negligible but scientifically important as it is derived from melts generally unaffected by crystal fractionation. Such glasses represent the best samples of the lunar mantle. In comparison to mare basalts, the non-mare volcanic features are relatively insignificant. Most notably, some obviously volcanic domes can reach diameters of 20 km and heights of ~1000 m. Prominent examples are the Gruithuisen domes, whose shapes suggest that they consist of viscous lava, and the Mairan domes, which might be explosive in nature.

8.6.2 Mercury

Large volcanic plains related to impact basins and in high northern latitudes do not display broad, rifted rises, constructional landforms (e.g., shield volcanoes), or individual linear, leveed lava flows. They were probably emplaced by flood-lava eruptions, and are not the products of narrow, leveed flows sourced from small dikes and more limited-volume surface eruptions. The volume of the northern volcanic plains is estimated to be at least 4×10^6 km^3 to 10^7 km^3, consistent with high-volume, high effusion-rate eruptions. Such volume is comparable to that of terrestrial LIPs (Columbia River Basalts: 1.3×10^6 km^3; Deccan Traps: 8.2×10^6 km^3; Greater Ontong Java igneous province: $\sim 7 \times 10^7$ km^3), and indeed the northern volcanic plains on Mercury are considered to be a planetary analogon to terrestrial LIPs.

Mercury was thought to be volatile-poor, hence no signs of explosive volcanic activity were expected prior to recent spacecraft exploration with high-resolution cameras. However, dozens of deposits have now been identified that are morphologically consistent with emplacement by explosive activity.

8.6.3 Venus

As for the tectonic landforms, the larger bodies (Venus and Mars) display a higher abundance, morphological diversity, and age range of volcanic landforms than the smaller bodies, Mercury and the Moon. Again, this is not surprising, as volcanism is linked to the endogenic activity of a planet, which in turn is coupled to its mass and tectonic evolution. Venus displays large volcanic provinces with huge individual edifices (shield volcanoes) that can reach diameters of more than 1000 km, and extensive lava flow fields that can cover plains with a lengths of ~1000 km. Rectilinear to sinuous *channels* within Venusian lava plains can reach extreme lengths of ~6800 km (Nile river: ~6500 km). Lavas with extremely low viscosities are required to build channels of such lengths, and basaltic lavas typically do not have such low viscosities. Other compositions like ultramafic komatiites, high-Ti lunar-type basalts, and carbonatites or sulphur flows may be candidate alternatives. Venus also shows many small individual volcanic cones or shields, which form volcanic shield fields. The high atmospheric surface pressure is thought to prevent rapid and violent magma degassing, analogous to the high water pressure on Earth's ocean floors. Some researchers have suggested that submarine volcanic landforms may indeed be useful analogues for the study of volcanism on Venus. Although some landforms have been tentatively interpreted as products of explosive eruptions, including a pyroclastic flow deposit, no conclusive morphologic evidence of explosive eruptions has been observed on Venus. On the other hand, there are several very large surface features on Venus that do not have terrestrial analogues. They appear to have a volcano-tectonic origin and have been classified as magmatic centers. In many cases, they have axisymmetric planform geometry and are characterized more by structural features indicating surface deformation than by magmatic features like volcanic edifices or lava flows. This class of features includes *coronae*, *arachnoids*, and radial fracture centers, or *novae*. Coronae can have diameters of up to 2000 km and are characterized by a concentric ring of tectonic fractures that surround other fractures and volcanic landforms such as domes and flows. Coronae as well as other large magmatic centers are thought to be related to mantle plumes (upwelling), but downwelling models have also been proposed. Recent thermal observations from orbit were interpreted as indicative of active volcanism, but due to the lack of modern high-resolution radar imaging it is not possible to search for associated landforms.

The distribution of volcanoes suggests that magmatism and internal dynamics on Venus are driven by large-scale mantle convection and mantle plumes, based on similarities with the distribution pattern of hot-spot volcanism on Earth. Indeed, Venus may be considered an observable analogue for the Archean Earth before the onset of plate tectonics (Chap. 11).

8.6.4 Mars

The surface of Mars shows numerous volcanic landforms of diverse morphology, which were produced by both effusive and explosive eruptions. The best-known examples are huge basaltic shield volcanoes. The largest of them, Olympus Mons, has a basal diameter of ~600 km, a height of more than 20 km, and is probably the largest volcano in the Solar System. These edifices share many physiological characteristics with large terrestrial shields: very low flank slopes (~5°), and numerous large individual lava flows that can be channel-fed or tube-fed. It seems that these volcanoes were built over a very long time, perhaps billions of years. Huge calderas formed in several stages, and show that large shallow magma chambers must have existed beneath their surface (Fig. 8.7). The youngest lava flows on their surfaces are only a few tens of million years old; some may even be as young as a few million years. Lava flows on Mars can reach extraordinary lengths of several hundred kilometres, indicating high volumetric flow rates for extended periods of time. Lava flows of this length are rarely, if ever, observed on the Earth, and it is debated whether their formation on Mars requires some extraordinary conditions. The rheology of Martian lava flows, as determined from their morphology and morphometry in remote sensing data, suggests basaltic lavas with low viscosities.

Hundreds of much smaller shield volcanoes are distributed in clusters within the Tharsis and Elysium provinces (Fig. 8.14). Their diameters range from several to tens of kilometers, their heights reach a few hundred meters, and the corresponding flank slopes are extremely shallow (~0.5°).

Since methane in Mars' atmosphere was first reported on the basis of orbital and Earth-based observations back in 2004, and later confirmed by MSL rover measurements, researchers tried to explain how it is currently supplied to the atmosphere where its survival time is short. Since a biologic origin seems unlikely, possible sources may involve geologic processes such as serpentinization. On Earth, methane is a common product of mud volcanism (sometimes also called sedimentary volcanism). Therefore, several types of pitted cones have been interpreted as mud volcanoes in the wake of the possible methane detections. As the shape of mud volcanoes can resemble that of igneous volcanoes (e.g., monogenetic scoria cones or tuff cones), interpretations can be ambiguous, and in several cases both origins (igneous and sedimentary) seem possible.

8.6.5 Io

As a consequence of tidal heating by Jupiter, Io experiences a very high surface heat flow and a volcanic activity that surpasses that of the Earth. Volcanic landforms are abundant and can be classified into a few groups: Volcanoes, pyroclastic materials, and mountains that are probably no active volcanoes, but have a volcano-tectonic origin. Calderas on Io tend to be very large (average diameter ~40 km, maximum

diameter 200 km), hence this may hint at larger sizes of magma chambers than on Earth. Some very large lava flows on Io are considered to be analogues to terrestrial flood basalt fields.

8.6.6 Icy Bodies

Although unambiguous morphological evidence for cryovolcanism is rare on icy bodies in the outer Solar System, it has been demonstrated that it is an active process on Enceladus and Triton and was possibly active on Europa, Titan and Kuiper Belt objects such as Pluto and Charon. One definition of cryovolcanism is the extrusion of liquid water and gases that would be frozen solid at the ambient surface temperatures of the respective body. The energy to drive cryovolcanism is thought to come from tidal forces. Cryovolcanism is very important for the emerging scientific discipline of astrobiology, as it is a form of exchange between subsurface oceans that are thought to underlie the icy shells of some satellites, and the surface and space. As such, it delivers plumes to space that could potentially contain volatile species and organic compounds that may be biosignatures. Sampling such cryovolcanic plumes by spacecraft would then be an efficient method to analyse such materials, whereas the penetration of icy planetary shells by, e.g., melting probes may technologically not be feasible for some time to come.

Take-Home Messages

Most, but not all volcanic landforms on Venus, Mars, Mercury, and the Moon have morphological analogues on the Earth, and seem to be predominantly basaltic in composition.

The larger terrestrial planets do have a greater diversity in volcanic morphology than the smaller ones, and the presence of water/ice on Mars makes it the most diverse besides the Earth.

Differences in morphometry, shape, and other properties (e.g., particle size distribution), may be explained by the different environmental conditions, e.g., the atmospheric pressure and the gravitational acceleration.

There are only a few tectonic landforms on the terrestrial planets (e.g., the coronae on Venus) that do not have direct counterparts on the Earth.

It is expected that the continuing exploration of the Solar System with spatially higher-resolving instruments in the future (e.g., better radar imaging of Venus) will enable identifying new and smaller-scale endogenic landforms.

Of special interest are small rocky-icy bodies such as Europa, Enceladus, Titan, Triton, and Kuiper Belt objects such as Pluto and Charon, as they seem to be much more geologically active than expected.

Suggested Readings

Ernst, R., Grosfils, E., Mège, D.: Giant dike swarms: Earth, Venus, and Mars. Annu. Rev. Earth Planet. Sci. **29**(1), 489–534 (2001). doi:10.1146/annurev.earth.29.1.489

Grott, M., Baratoux, D., Hauber, E., Sautter, V., Mustard, J., Gasnault, O., Ruff, S.W., Karato, S.-I., Debaille, V., Knapmeyer, M., et al.: Long-term evolution of the Martian crust-mantle system. Space Sci. Rev. **174**(1–4), 49–111 (2013). doi:10.1007/s11214-012-9948-3

Harris, L.B., Bédard, J.H.: Crustal evolution and deformation in a non-plate-tectonic Archaean Earth: comparisons with Venus. In: Evolution of Archean Crust and Early Life, pp. 215–291. Springer, Berlin (2014). doi:10.1007/978-94-007-7615-9_9

Hartmann, W.K.: The giant impact hypothesis: past, present (and future?). Philos. Trans. R. Soc. **372**, 20130249 (2014). doi:10.1098/rsta.2013.0249

Hauber, E., Wagner, R.: Planetary geology: craters and chronology, volcanism, tectonics. In: Trümper, J. (ed.) Solar System, pp. 345–433. Springer, Berlin, Heidelberg (2009)

Head, J.W., Wilson, L.: Generation, ascent and eruption of magma on the Moon: new insights into source depths, magma supply, intrusions and effusive/explosive eruptions (Part 2: Predicted emplacement processes and observations). Icarus **283**, 176–223 (2016). doi:10.1016/j.icarus.2016.05.031

Lopes, R.M.C.: Chapter 35, Io: the volcanic moon. In: Breuer, D., Johnson, T.V.B.T. (eds.) Encyclopedia of the Solar System, pp. 779–792. Elsevier, Boston (2014). doi:10.1016/B978-0-12-415845-0.00035-9

Marov, M.Y., Grinspoon, D.H.: The Planet Venus. Yale University Press, New Haven (1998)

Mège, D.: Uniformitarian plume tectonics: the post-Archean Earth and Mars. Geol. Soc. Am. Spec. Pap. **352**, 141–164 (2001). doi:10.1130/0-8137-2352-3.141

Melosh, H.J.: Global tectonics of a despun planet. Icarus **31**(2), 221–243 (1977). doi:10.1016/0019-1035(77)90035-5

Platz, T., Massironi, M., Byrne, P.K., Hiesinger, H. (eds.): Volcanism and Tectonism Across the Inner Solar System. Lyell Collection, Special Publications, vol. 401, 448 pp. The Geological Society of London, London (2015). doi:10.1144/SP401.22

Schmincke, H.-U.: Volcanism, vol. 28, 324 pp. Springer Science & Business Media, Berlin (2004)

Schultz, R.A., Hauber, E., Kattenhorn, S.A., Okubo, C.H., Watters, T.R.: Interpretation and analysis of planetary structures. J. Struct. Geol. **32**(6), 855–875 (2010) doi:10.1016/j.jsg.2009.09.005

Solomon, S.C.: On volcanism and thermal tectonics on one-plate planets. Geophys. Res. Lett. **5**(6), 461–464 (1978). doi:10.1029/GL005i006p00461

Tanaka, K.L., Schaber, G.G., Chapman, M.G., Stofan, E.R., Campbell, D.B., Davis, P.A., Guest, J.E., McGill, G.E., Rogers, P.G., Saunders, R.S.: The Venus geologic mappers' handbook. Technical Report, US Geological Survey (1993). http://pubs.er.usgs.gov/publication/ofr93516

Watters, T.R., Schultz, R.A.: Planetary Tectonics, vol. 11, 480 pp. Cambridge University Press, Cambridge (2010)

Wilson, L.: Chapter 5 - Planetary Volcanism. In: Spohn, T., Breuer, D., Johnson, T.V. (eds.) Encyclopedia of the Solar System, 3rd edn., pp. 101–119. Elsevier, Boston (2014). doi:10.1016/B978-0-12-415845-0.00005-0

Wilson, L., Head, J.W.: A comparison of volcanic eruption processes on Earth, Moon, Mars, Io and Venus. Nature **302**(5910), 663–669 (1983). doi:10.1038/302663a0

Wilson, L., Head, J.W.: Mars: review and analysis of volcanic eruption theory and relationships to observed landforms. Rev. Geophys. **32**, 221–263 (1994). doi:10.1029/94RG01113

Zimbelman, J.R., Gregg, T.K.: Environmental Effects on Volcanic Eruptions: From Deep Oceans to Deep Space, 206 pp. Springer Science & Business Media, Berlin (2000)

Chapter 9
Surface Processes

Nicolas Mangold

9.1 Introduction

Planetary surfaces are shaped by a large variety of processes. The presence of an atmosphere is certainly the most important parameter that controls these processes. For this reason, Mars and Saturn's satellite Titan are the planetary surfaces that have the largest diversity of landforms and the closest resemblances to Earth. Nevertheless, the lack of an atmosphere does not mean the lack of surface processes. For instance, mass wasting and sublimation exist without the presence of an atmosphere. However, the surfaces of atmosphere-free bodies are more directly controlled by endogenous processes and impact craters compared to atmosphere-bearing bodies, explaining that this chapter focuses more on the latter. Gravity, temperatures and composition—especially the presence of volatiles—are also major parameters controlling surface processes. There is a huge amount of literature describing surface processes on Earth. The goal of this chapter is not to provide an exhaustive summary but to focus on surface processes that are most relevant for planetary bodies, taking Earth as reference and applying the knowledge acquired on Earth to planetary bodies.

N. Mangold (✉)
Nicolas Mangold, Laboratoire Planétologie et Géodynamique, CNRS and Université de Nantes, UMR 6112, Nantes, France
e-mail: Nicolas.Mangold@univ-nantes.fr

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_9

9.2 Eolian Transport and Erosion

9.2.1 Entrainment of Grains by Wind

Most of the eolian processes are controlled by relatively simple physical laws in which wind is the central agent. Friction velocity corresponds to the velocity of the wind at the contact with the surface. Wind has the capability of dislodging and entraining grains from the ground into the air. The friction velocity u_* depends on τ_0, the shear force on the ground, and, ρ, the density of the atmosphere:

$$u_* = \sqrt{\frac{\tau_0}{\rho}} \tag{9.1}$$

The wind velocity U increases logarithmically with height above the surface (Fig. 9.1a), and can be defined as a function of height z:

$$U = \ln\left(\frac{z}{z_0}\right) \cdot \frac{u_*}{k} \tag{9.2}$$

where k is a constant usually taken as 0.4 and z_0 is the roughness length. The roughness of the surface is important because it will create variations in the local friction enhancing entrainment of particles located on the top of local heights. On sand sheets, roughness can be related to grain size only, or to bedform wavelengths.

The threshold velocity u_{*_t} is defined as the velocity at which grains start to move. It is common to use a static and a dynamic threshold. Indeed, after the movement of

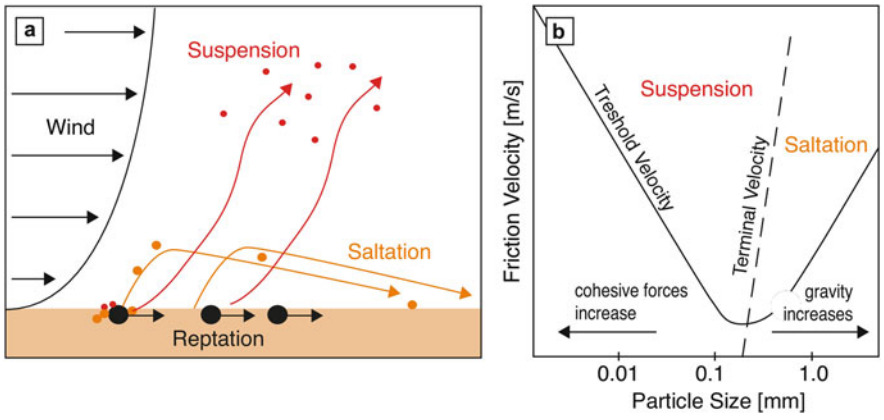

Fig. 9.1 (**a**) Sketch of basic wind entrainment processes. Wind decreases in intensity close to the ground. Reptation moves large grains on the ground. Saltation moves sand-size grains up into air before they fall down. Suspension moves dust into the atmosphere; (**b**) Ideal profile of threshold friction velocity for various grain sizes. The minimum of the curve indicates the first grain size to move when wind reaches the threshold velocity

grains is generated, the velocity necessary to move the disturbed grains is lower. The threshold friction velocity depends on two main forces that have inverse effects. For small grains, cohesion is predominant. Cohesion is directly linked to electrostatic forces, which is a function of $(1/d^3)$ with d being the diameter of particles. For large grains, gravity is predominant, which increases progressively the threshold velocity needed for lifting grains (Fig. 9.1a). The combined result of the two opposite forces has a minimum, which means that wind will move first grains in the range of this minimum. Experimental studies have shown that this minimum varies with parameters such as grain composition, moisture or grain shapes, but it always corresponds to diameters in the 50–500 μm range (Fig. 9.1b).

After entrainment of grains has started, it has been observed in experiments and in the field that grains behave also differently depending on their size. The largest grains (granules, > 1 mm diameter) move by rolling or sliding on the ground; this movement is called *reptation* (or sometimes creep). Sand-size particles are also lifted in the air and are temporally entrained by the wind, but gravity tends to bring them back to the ground. This process is called *saltation*, because sand particles usually fall, impact other grains and are transported again a few steps farther away, and so on. The effect of saltation is strong on other grains, so that dust particles, not directly entrained by wind due to their higher cohesion, are dislodged from the surface and, then, entrained by wind. Dust lifted in the air forms clouds of particles in suspension (Fig. 9.2a). Suspension is strongly dependent on the turbulence of the wind whereas saltating sand grains are too large to be affected. Suspension can be short-term for large dust particles that settle down after the wind stops, or long-term for finest dust particles transported by global wind circulation.

This process is similar on all planetary bodies with an atmosphere, with strong variations related to atmosphere density and gravity, and, at a lower degree, grains composition and density. Typical terrestrial (static) threshold velocities are ca. $0.2\,\mathrm{m\,s^{-1}}$. Remember that these values are friction velocities at the surface while wind is actually measured at 1 m, or more, above the surface. Terrestrial field observations have shown that these friction velocities translate into critical wind speed $>5\,\mathrm{m\,s^{-1}}$ at 1 m above ground. The lower pressure of the atmosphere on Mars requires threshold velocities that are higher by one order of magnitude than on Earth, ca. $2\,\mathrm{m\,s^{-1}}$, requiring stronger winds as well. Titan and Venus have higher atmospheric densities enabling lower thresholds of approximately 0.04 and $0.02\,\mathrm{m\,s^{-1}}$, respectively. In contrast to these strong variations in threshold velocities, the minimum grain size for which this critical threshold is measured remains relatively similar from body to body, and is in the diameter range of 100–400 μm.

9.2.2 Dunes and Eolian Sandstones

The formation of dunes has occupied researchers for decades and we only provide here a short summary. Entrainment by wind has direct implications on the building

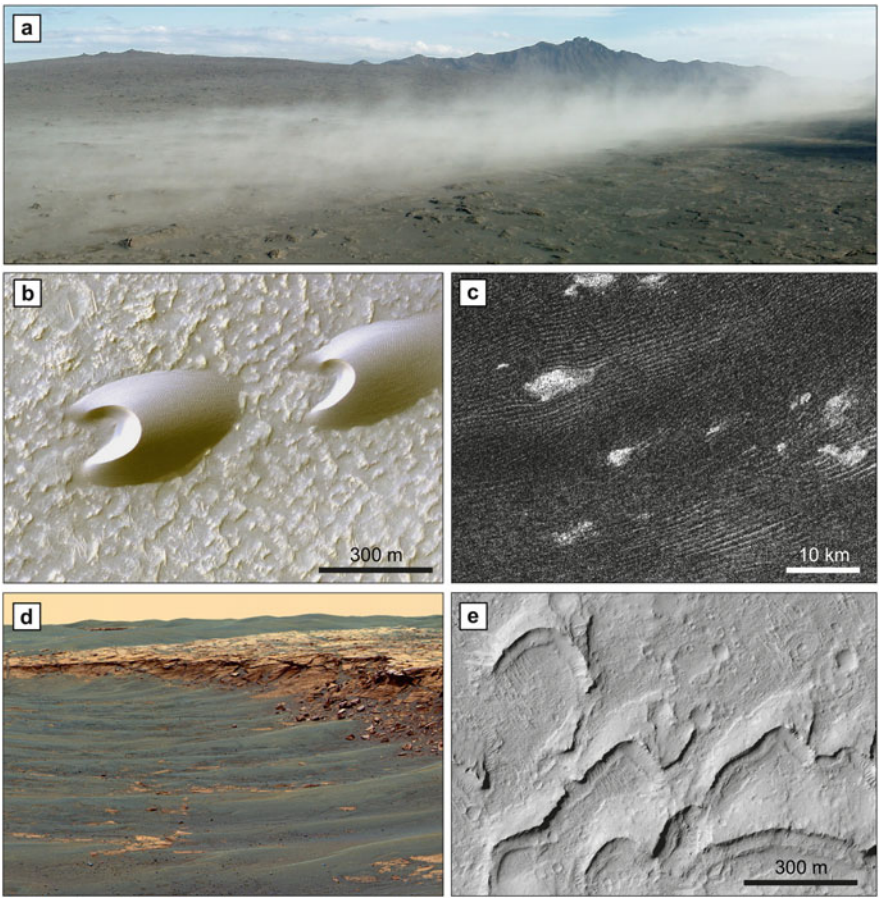

Fig. 9.2 Eolian features on the Earth, Titan and Mars. (**a**) Dust lifted in the air by saltating grains after wind reached threshold velocities in Iceland (wind direction from left to right); (**b**) barchan dunes on Mars. Note the wind is from right to left, as indicated by the position of the long side of the dunes; (**c**) longitudinal dunes on Titan; (**d**) ripples on Meridiani Planum; (**e**) cemented barchan dunes recognizable by their crescent shapes. Boulders, sharp crests and texture indicate these dunes are not currently forming. Source: (**a**) N. Mangold. (**b**) and (**e**) NASA MRO HiRISE Team. (**c**) NASA/JPL/Cassini Team. (**d**) NASA/JPL

of sand dunes. First, sorting of grains by diameter occur for sand grains through the different processes of reptation, saltation and suspension. Most of the dunes form transverse to the main wind direction. The segregation of patches of sand represents a roughness against wind, enhancing the entrainment on the windward side. In contrast, on the lee side, particles are not moved by wind anymore. Dunes become asymmetric, with a gentle slope on the windward side, and a steep slope on the lee side. Then, the lee side becomes controlled by gravity flows when exceeding the angle of repose of the particles forming the dune. At this point, the dune can

start to move in the wind direction, controlled by accumulation on the windward side, and a slipface on the lee side. Migration of dunes can reach several meters by year on Earth.

Although the process is relatively universal, the various wind regimes can create a variety of dune shapes. *Barchans* are crescent-shaped dunes that formed orthogonally to a dominant wind direction. The long side of barchans indicates the direction from where the wind comes, whereas the small side is made of a steeper slipface (Fig. 9.2b). The crescent shape is controlled by the fact that the thinner sections of the dunes on both sides migrate faster than the thick central section. *Star dunes* are shaped by multiple wind directions forming usually huge dunes that do not migrate much. Linear dunes are parallel dune fields that can form by two types of wind regime. Longitudinal linear dunes (also called *seif* form parallel to the main wind direction, like initial barchan dunes becoming elongated by strong winds. These dunes form large dune fields in the Sahara and in the equatorial regions of Titan (Fig. 9.2c). The northern plains of Mars also display widespread areas with longitudinal dunes. Linear dunes can also be oriented transverse to the main wind direction as observed for instance on the floor of Martian valleys where the wind flows along the valley. These transverse dunes can usually be distinguished from longitudinal dunes by their undulated shape, the asymmetric cross section, and the local presence of barchans, i.e. barchanoids. Mars displays an intriguing type of *transverse eolian ridges (TARs)* corresponding to a particular type of small transverse dunes.

Dune is the word used for a sand body in the 1 m–1 km scale. Below 1 m, bedforms are named ripples. At the surface of sand sheets, ripples develop orthogonally to the main wind direction as a consequence of impacts due to saltation. Such ripples are observed on the surface of dunes on Mars. Nevertheless, ripples are also observed outside of sand sheets, developing on bedrock, showing coarser grains (mm-size granules) that accumulate towards the crest of ripples in contrast to saltation effects sorting smaller material at the top. In this case, ripples are due to reptation of coarse sand grains and granules (mm-size grains). To distinguish them from sand ripples on sand dunes, these ripples are frequently named granule ripples, as observed on Mars at Meridiani Planum (Fig. 9.2d). By extension, large ripples reaching the size of dunes are sometimes called megaripples.

The composition of sand in dunes is affected by sorting effects caused by variations in density, hardness or shape of grains and chemical stability to alteration. Earth is dominated by quartz dunes because quartz is frequent and is not easy to alter chemically. Locally, gypsum dunes or basaltic dunes are observed in regions with, respectively, evaporitic playa deposits and volcanic bedrock. Martian dunes are mostly dark displaying a basaltic composition dominated by pyroxene and olivine minerals. Dunes on Titan were initially interpreted as water-ice grains, but recent findings consider organic particles to be equally possible, either from atmospheric flocculation or reworking of lacustrine deposits.

When accumulated sand sheets may start to be cemented in sandstones. Aeolianite is a term often used for sandstones formed from wind-blown material. Eolian sandstone structure often includes cross-beddings with significant angles

(up to ∼30°) as a consequence of the internal architecture of sand dunes. Eolian sandstones usually do not preserve the original dune morphology, but examples of in place cemented dunes exist on Mars where paleodunes were identified by the presence of boulders and their rough texture compared to smooth sand dunes (Fig. 9.2e). These barchans were likely cemented by processes involving local moisture or water circulation, frost, or salts, and were never buried deep beneath other sediments, preserving the dune morphology. Sedimentary rocks showing cross-beddings formed by eolian deposition have been observed by the Opportunity rover at Meridiani Planum and the Curiosity rover at Gale crater. At Meridiani Planum, eolian sandstones have been interpreted as a cementation of basaltic sand dunes by salt-rich groundwater circulation.

9.2.3 Loess, Dust and Duststones

Wind entrainment and has a direct implication on lifting dust in suspension, that otherwise would be stuck on sand grains. On Earth, this process creates, for instance, net flux of dust out of the Sahara towards the Atlantic Ocean. This process was especially efficient during ice ages as a consequence of glaciers crushing bedrock in small particles, from which the finest grains were transported by wind subsequently. Terrestrial dust is usually defined by clay-size particles, smaller than 10 μm, and is composed by particles of varied composition, including secondary minerals such as clay minerals. On Mars, dust particles are <1 μm due to the thin atmosphere. Martian dust contains basaltic minerals and is reddish due to the presence of iron oxides, giving to the Martian surface its famous red color.

Dust devils are spectacular features formed by a specific wind entrainment. They form by a vortex created by thermal flux from surface heating by the sun. They usually form in the afternoon, after the sun has heated the ground, as observed both on Earth and Mars. They rotate along a vertical axis and display also a movement of translation. Mars mission Spirit has observed many dust devils in Gusev crater plains (Fig. 9.3a). Many images of the Mars surface show intricate dark streaks with local spirals corresponding to dust devil tracks. The dark albedo comes from the removal of dust, which is especially visible on sand dunes covered by a thin dust mantling (Fig. 9.3b). Dust devils may be the major process lifting dust on Mars. Large dust storms can fully hide the surface of Mars. The transition from dust devils or local sand storms to regional and global storms is not fully understood.

On Earth, loess deposits are related to the transport and deposition of wind-blown material through atmospheric circulation. Loess deposits are famous in northern China where hundreds of meters of deposits were sedimented during past ice ages. Mars displays also thick dust deposits. For instance, the region of Arabia Terra plays the role of regional sink for dust; with accumulation of dust reaching 10–100 m after hundreds of millions of years of accumulation. It is difficult to reactivate dust deposits because of the cohesion of grains of that size, but it has been show that dust form aggregates that may be able to be entrained by wind by saltation,

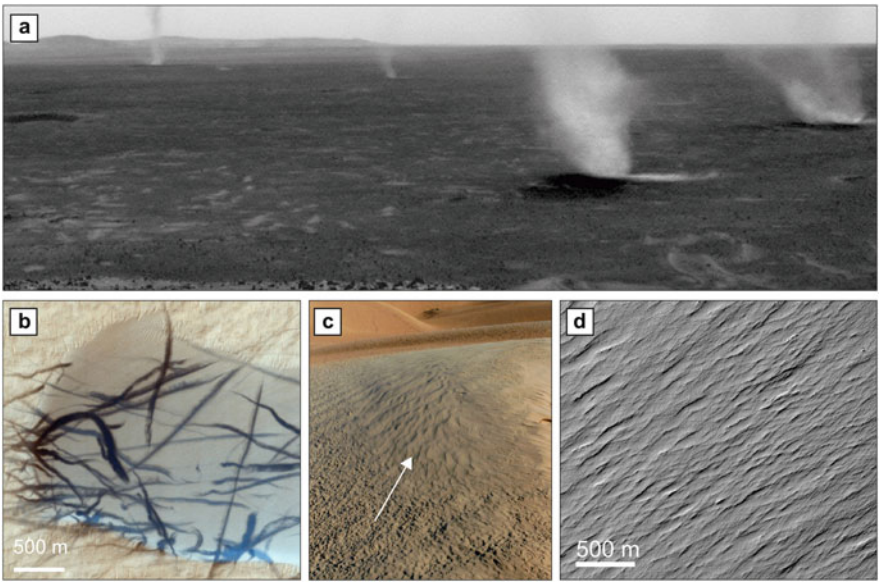

Fig. 9.3 Eolian features on the Earth and Mars. (**a**) Dust devils as seen by MER rover Spirit at Gusev Crater; (**b**) dust devil tracks revealing darker sand; (**c**) patch of silt size material above a dune in Tunisia; note the sinuous shapes formed by the erosion by wind; (**d**) sinuous grooves carved by wind erosion on fine-grained deposits at Medusae Fossae Formation. Source: (**a**) NASA/JPL. (**b**) NASA MRO HiRISE Team. (**c**) N. Mangold. (**d**) ESA Mars Express HRSC

before breaking into smaller particles. Specific erosional wind patterns form on dust deposits such as sinuous elongated flutes (Fig. 9.3c, d).

9.2.4 Wind-Related Patterns

Wind can sculpt rocks in various patterns if the surface is stable long enough for such an abrasion. Rocks modified by wind abrasion are generically named ventifacts and can include three-faceted rocks, pits, flutes or etched rocks. The main macro-scale landforms identified on orbital image consists of yardangs, a Turkish word meaning steep. Yardangs are observed on Earth when sustained mono-directional winds sculpt elongated (up to 1:10 of width:length ratio) hills inside soft material, usually lacustrine or fine-grained eolian sediments (Fig. 9.4a). Yardangs are observed at all scales, from few centimeters to kilometers, and require thousands of years to form. Yardangs on Mars are frequent, especially in regions of fine-grained deposits exposed to strong wind such as Medusae Fossae Formation (Fig. 9.4b).

Wind can create wind streaks in presence of topographic obstacles. Wind streaks can form either by deposition or erosion of material. A wind streak can form when material is transported from an butte, like a volcano, depositing grain of different

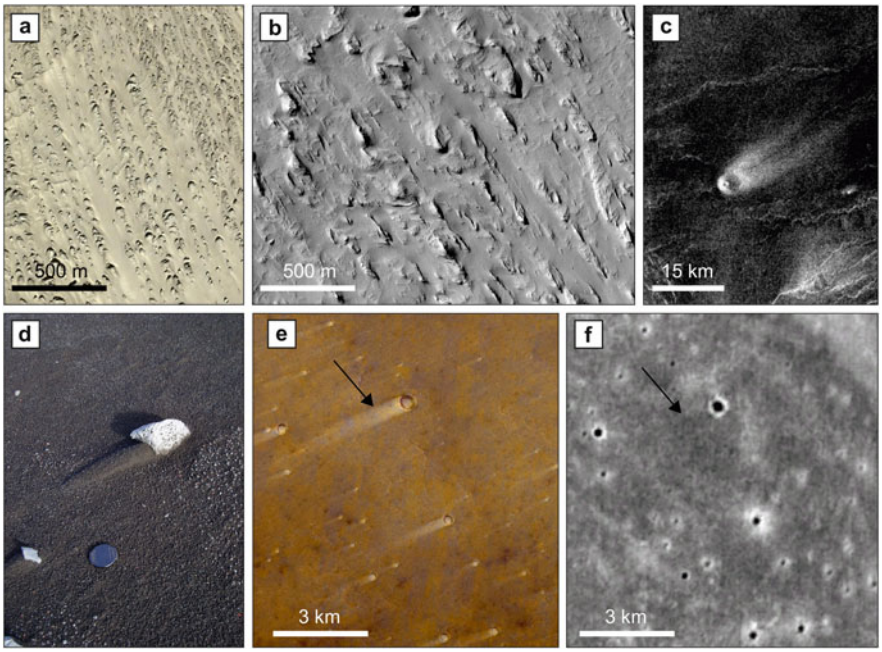

Fig. 9.4 Eolian features on the Earth, Venus and Mars. (**a**) Yardangs in China, Takla-Makan desert; (**b**) yardangs in Medusae Fossae Formation, Mars; (**c**) wind streak behind a volcano on Venus surface; difference in tone represents difference in material roughness in radar data; (**d**) small wind streaks made of sand preserved from erosion by a stone. This example illustrates in Iceland that streaks can be due to differential erosion; (**e**) wind streak in Syrtis Major Planum, Mars; light color indicates more dust preserved behind the crater; (**f**) THEMIS mosaic of the same area lacking wind streaks, demonstrating that streaks are thin landforms. Source: (**a**) Digital Globe. (**b**) NASA MRO HiRISE Team. (**c**) NASA/Magellan. (**d**) N. Mangold. (**e**) ESA Mars Express HRSC. (**f**) NASA/THEMIS/ASU

albedo than those of surrounding plains; i.e. depositional wind streak. But a wind streak can also form when the topography protects the shadow behind it from the accumulation of eolian material everywhere else, i.e. differential deposition streaks, as observed on Venus (Fig. 9.4c). This process does not require the volcano to be active. Ash deposition due to the volcanic activity itself can be distinguished by an aureole larger than the shadow from the volcano. Wind streaks can form also by erosion of material deposited. Obstacles such as rock outcrops or impact craters play the role of shadows protecting from erosion the deposits in the lee of wind (Fig. 9.4d). Wind streaks of that category are frequent on Mars. They are light-toned as due to residual dust deposits preserved from erosion by wind. Wind streaks can reach tens of kilometers in length but are usually thin. The lack of streaks on thermal images of the same area shows that those streaks do not exceed a few centimeters in thickness (Fig. 9.4e, f).

9.3 Fluvial Erosion and Deposition

9.3.1 Rivers: Diluted Flows

Fluvial erosion is ubiquitous on Earth but is also one of the most complex processes, explaining that a large community is still working on its basic principles. The channel is defined as the location where the river flows, whereas the fluvial valley is the elongated depression formed by the prolonged action of rivers. A fluvial valley is usually much larger than the channel(s) that flows on its floor. A fluvial basin is usually composed of valleys incising into bedrock in the upstream section of the basin, alluvial plains with fluvial sediment accumulations and depositional fans at the outlet (Fig. 9.5a).

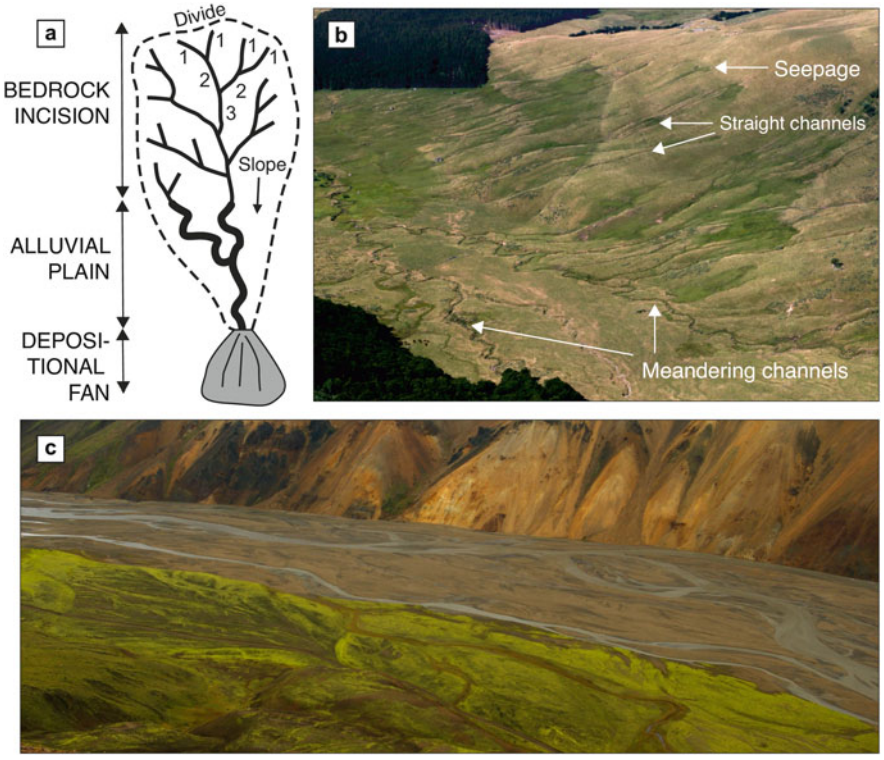

Fig. 9.5 Fluvial systems on the Earth. (**a**) Sketch of a drainage basin with three main sections (erosion of bedrock, alluvial plains and terminal deposition); numbers indicate the order of tributaries in Horton–Strahler distribution; (**b**) image within a catchment in Auvergne, France; straight streams on steep slopes connect to meandering flows on gentler slopes; seepage at valley head is close to the basin boundary showing limited groundwater contribution; (**c**) braided rivers in Iceland. Source: (**a**), (**b**) N. Mangold

We first focus on the hydrological part of the process, i.e. channel dynamics. The mathematics of channel hydrology is relatively well-understood although several parameters are not always known. Dilute river flows have a Newtonian behavior: A linear relation links the shear stress τ to the strain γ rate through the fluid viscosity μ:

$$\tau = \mu \frac{\mathrm{d}\gamma}{\mathrm{d}t} \tag{9.3}$$

The flow is known as laminar when the velocity gradient decreases with depth (similar as wind, Fig. 9.1b). Velocity is strongly linked to flow resistance that is related to the friction at the base of the channel. This law is, however, not applicable directly to fluids that are turbulent as are most of the rivers and more complex laws have been used. Hydrological laws were initially defined by empirical properties assuming an average velocity V. One of the first law proposed is the Manning equation that gives the discharge rate Q as:

$$Q = A \cdot V \tag{9.4}$$

$$= A\sqrt{S \cdot \frac{R^{4/3}}{n^2}} \tag{9.5}$$

where A is the cross-sectional area, R is the hydraulic radius (the ratio of flow cross sectional area A to wetted perimeter P), S is the channel gradient, and n is the Manning coefficient. This coefficient is defined according to the basal roughness defined by the proportion of pebbles (from 0.01 when smooth to 0.1 when very rough). This law has been modified for planetary applications to:

$$Q = A\sqrt{S \cdot g_p \cdot \frac{R^{4/3}}{g \cdot n^2}} \tag{9.6}$$

where g_p is the gravity of the planet considered and g is the terrestrial gravity. The weakness of this law is that the Manning coefficient is not easy to quantify. The Darcy–Weisbach equation is now more often used and recommended:

$$Q = A\sqrt{\frac{8g_p \cdot R \cdot S}{f}} \tag{9.7}$$

where f is the friction factor that depends on channel-bed clast distribution. The friction f is derived empirically for gravel bed regime by the relation:

$$f = \frac{8}{5.75 \log_{10}\left(\frac{R}{d_{84}}\right) + 3.514} \tag{9.8}$$

where d_{84} is the size at which 84% of clasts are that size or smaller, which is measurable for terrestrial rivers. This parameter d_{84} is still a problem for planetary applications for which there are poor constraints on clast size distribution. Another problem for quantifying channel discharge rates is the fact that planetary surfaces show paleo-channels no more fed by water, as typically is the case on Mars. There, only bankfull discharge rates can be deduced from topographic observations of empty paleochannels and those values are often overestimated.

Channels display various shapes depending on slopes and discharge rates (Fig. 9.5b, c). Channels are straight only on steep slopes ($>1°$) and limited discharge rates ($<100\,m^3\,s^{-1}$). Channels display *meanders* where the slope is typically ca. 0.5–1° but this behavior depends on discharge rates and solid content. *Braided* channels consist of series of channels separated by fluvial bars. Channels cross themselves, migrate with time, and can become a single channel at high water level (bankfull discharge). Braided streams form for higher flow dynamics than meandering rivers, thus for combined effects of relatively higher slopes and/or higher discharge rates. *Anastomosed* is a term used as a synonym for braided, although this confusion should be avoided. Anastomosed channels display crossing channels similar in planview to braiding but without migration, i.e. channels stay year after year at the same location. Anastomosed channels exist mainly for very flat areas covered by vegetation. Except for some of the latter anastomosed channels, none of the channel types are stable in time. They can evolve from one to another types depending on variations in solid content and discharge rates.

Fluvial valleys are formed by the action of rivers over time, but bedrock erosion is not limited to the erosion by water flowing in the channel. Mass wasting on valley sides, presence of soils/gravels at the surface, or water infiltration are fundamental processes to take into account as well. Two approaches dominate the study of valley networks: a geometric approach, describing the basin geometry/topology, and a physical approach describing erosion processes complementary to the channel hydrology analyses. We mainly detail hereafter the first approach that is easy to use on planetary bodies with available imagery and topography.

The topology of valley networks has been studied for a century or so. Horton and Strahler, in the mid-twentieth century, have introduced a way to classify the distribution of valleys. The smallest tributaries are of order 1. An order 2 is found when two tributaries of order 1 connect, and so on (Fig. 9.5a). Although there are other ways to classify tributaries (e.g., Shreve orders), the Horton–Strahler classification is the most common. In this classification, the stream lengths and drainage areas increase with the increase of orders; whereas the number of streams decreases (order 1 tributaries are always the most numerous). The meaning of order 1 tributaries is different in planetary applications than on Earth because orbital data cannot reach the level of precision of mapped terrestrial rivers, but this does not modify the validity of these laws.

The *drainage density* is a simple concept that enables to quantify the proportion of fluvial valleys in a drainage basin defined by the local topography. Drainage density (in m^{-1}) is the total length of all streams ratioed over the area of a basin. Applied to paleovalleys, it must be used with caution because ancient valleys have

been partly modified by resurfacing (e.g., volcanic flows), which could have altered the initial drainage patterns or topographic boundaries of former basins.

Various topological laws have been used on Earth for valley networks, among which Hack's law is certainly the most used. It has been found empirically by the measurements of a high number of drainage basins:

$$L = 1.4 \cdot A^h \tag{9.9}$$

Where L is the main stream length, A is the drainage area of the basin (upstream of the point it is measured) and h is the Hack exponent. It is observed empirically that h$\sim$0.57 on Earth in average (with notable variation around this value). This law has been followed up with fractal theories. If h =0.5, the drainage network is self-similar: the geometry would be similar at any scale. The average Hack exponent shows that terrestrial networks are close to self-similarity.

These various properties of drainage networks (tributaries hierarchy, connectivity, consistent Hack parameter, dense drainages) do not account directly for physical processes, but it is admitted from terrestrial observations that they can only be achieved if rivers are active simultaneously in a given basin (as formed by precipitation), and if they formed over a significant period of time (e.g., >1000 years) that enabled the development of this organization (Fig. 9.6a, b).

The transfer from precipitation to rivers is far to be a simple process. It varies with the presence of vegetation, of rocky debris or of a regolith, the amount of infiltration into bedrock, the evaporation of water at the surface and all hillslope processes in hilly areas. These processes are difficult to describe on planetary surfaces. A consequence has been the oversimplification of terms used on Earth for well-defined processes. For instance, the term *runoff* includes a number of different processes leading to transport of liquid water downslope, while the term *overland flow* is used for flows that occur on the ground after soils are saturated and infiltration becomes limited. Planetary scientists often use both terms indifferently indifferently to describe all water flows at the surface, in opposition to subsurface flows.

The term *seepage* is defined as the location where infiltrated water reaches the surface again. The action of overland flows and seepage is usually combined to develop the order-1 tributaries of drainage networks. An exception relevant for planetary applications is the infiltration of water into highly permeable bedrocks such as lava flows. In this case, water is able to infiltrate deep. The process of seepage at the base of steep canyon is defined as sapping and is usually invoked when deep canyons with amphitheater heads forms as a consequence of subsurface flows (T in Fig. 9.6c). Sapping-related valleys grow backward by regressive erosion. However, it has been shown that amphitheater-heads can also develop from overland flows, especially when the erosion is controlled by strong lithological variations. Such landforms should therefore not be used as a unique indicator of subsurface drainage.

Apart from Earth, Titan is the unique planetary body that displays active river flows, although these streams are filled by organic fluids—likely liquid methane—rather than liquid water (Fig. 9.6a). Some of these rivers join depressions that have

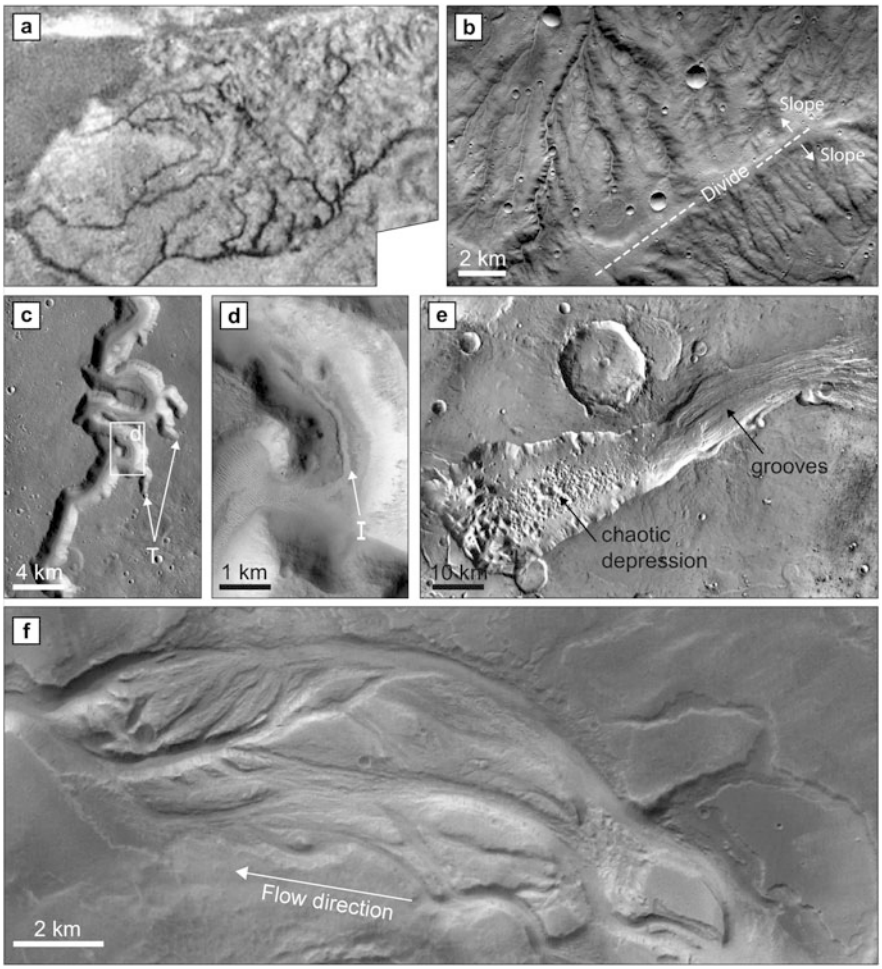

Fig. 9.6 Fluvial valleys on Titan and Mars. (**a**) Image of branched valleys on Titan; the lander Huygens has landed in the dry plain on the top left; (**b**) image of well-organized valley networks on martian highlands, note the topographic boundary which divides networks in two groups and the abundant gullying on the side of valleys showing runoff took place through all the area from precipitations (rainfall or snowmelt); (**c**) and (**d**) Nanedi Valles with 4 km wide meandering valleys bearing a 100 m wide residual channel on its floor; (**e**) Ravi Vallis, east of the Valles Marineris region is an outflow channel with a chaotic source region and grooved terrains indicating strong erosion; (**f**) close-up on Mangala Vallis, showing braided paleochannel with deep incision (10–100 m deep) showing intense erosion through large and deep channels. Source: (**a**) ESA Huygens. (**b**)–(**d**) NASA MRO/CTX/MSSS. (**e**) NASA/THEMIS/ASU. (**f**) ESA Mars Express HRSC

been shown to bear lakes. Individual valleys display locally *inner channels* showing that these valleys are not produced by a few flood events but by sustained activity. The geometry of Titan's rivers is similar to terrestrial rivers with a possible strong contribution of seepage, suggesting the presence of aquifers. Channels dynamics and sediment transport on Titan remain uncertain due to the lack of constraints on the particle nature and size, the magnitude of infiltration and of the fluid properties at −180 °C.

Mars displays a huge amount of ancient flow forms (Fig. 9.6b–d). Although not all are linked to sustained rivers, the presence of dense valley networks with a topology similar to terrestrial networks supports the past activity of sustained fluvial flows (Fig. 9.6b). The best organized valleys are dated of the Noachian period (∼3.6–3.8 Gyr). Hesperian valleys (∼3–3.6 Gyr) are also well-developed but more scattered. Amazonian valleys (<3 Gyr) are only poorly branched with low density and limited erosion, suggesting ephemeral activities. The mainstream length of Noachian and Hesperian valleys can reach several 100s of kilometers and depths of several 100s of meters. Drainage densities reach 0.1–1 km^{-1} in regions with well-preserved landforms, a magnitude similar to those of semi-arid regions of Earth. Local inner channels demonstrate that valleys formed by prolonged river activity. Inner channels were used to estimate bankfull discharge rates locally (Fig. 9.6d). Values from 10^2 to 10^4 m^3 s^{-1} are comparable with discharge rates measured in terrestrial fluvial activity.

Sapping processes related to subsurface flows were proposed to explain poorly-branched valleys with theater-heads such as Nanedi or Nirgal Valles (Fig. 9.6c). This process was especially invoked to explain fluvial valleys by permafrost melting in a cold climate without a role of surface water. However, these valleys display clear inner channels showing surface flows as well as sinuosities inconsistent with subsurface flows only (Fig. 9.6c–d). It is now admitted that variations in lithology, permeability and slope are responsible of these variations in drainage networks shape. Elongated sinuous valleys mimicking fluvial landforms can also be observed on the Earth's moon, Mercury and Venus, but those are linked to lava channels only. Mars displays locally lava channels as well, but their characteristics (such as punctual head and lack of inner channels) help to differentiate them from valley networks. Nevertheless, one must be cautious when interpreting as fluvial any martian valley that is poorly branched.

9.3.2 Flood Systems: Concentrated Flows

When a river contains a high density of solid particles, it is defined as a dense or concentrated flow. There is no well-defined solid-to-water ratio at which a flow becomes dense, but it is usually considered to be >0.01. The solid-to-water ratio can reach up to 0.3–0.5 in very concentrated flows. In such flows, turbulence is still acting in spite of the high number of particle collisions. Above this solid:water ratio, the flow becomes more laminar and is better defined by debris flows with viscous

properties. It has been shown on terrestrial drainages that sustained, dilute flows ($\ll$0.01 of solid-to-water content) do not erode significantly the bedrock. Most of the erosion takes place during the few days per year when rivers carry a higher solid content, i.e. during flood events. Concentrated flows therefore play a strong role in the incision of valleys into bedrock and in the transport of large particles.

Concentrated flows are not always formed by flood events related to extreme climate episodes or sudden snowmelts. Concentrated flows can be triggered by punctual or low frequency phenomena such as lake overflows, natural dam failures, subglacial discharges or snow/ice- melting from volcanic activity. This last example has a dramatic illustration in Iceland where the so-called *jökulhlaups* are due to subglacial lake discharges in relation of volcanic eruptions. Jökulhlaups generate flows with a high particle content (0.1–0.5 rock-to-water ratio) and a high discharge rates ($>1.10^5\,m^3\,s^{-1}$, i.e. similar or higher than the Amazon river at its outlet). In general, concentrated flows are ephemeral and catastrophic, which is a strong contrast with the more gradual evolution of diluted flows. The understanding of these processes on Earth is biased by their low frequency at the scale of mankind evolution. Nevertheless, it has been shown from ancient channel beds in USA (in the Scablands of the Washington State) that the disruption of glacial lakes from the last ice ages generated locally huge floods with discharge rates higher by one order of magnitude than jökulhlaups.

Apart from Earth, only Mars displays such type of landforms. Catastrophic flood events on Mars were inferred from the observation of the so-called *outflow channels*. Several regions of Mars show large (1–100 km) and long (100–1000 km) valleys that dissect volcanic plains (Fig. 9.6e, f). The term *channel* is misused here as such large topographic depressions should be called valleys. But the reason of this choice was motivated by the observation of frequent braiding with fluvial bars, as well as tear-drop shaped islands, grooves and flute marks indicating strong bedrock erosion by a flowing fluid at a giant scale (Fig. 9.6f). A high solid content likely explains the erosional strength of these flows. Outflow channels have a poor organization (no hierarchy, localized valley heads) that differentiates them easily from the organization of valley networks. Although it has been proposed that outflow channels are turbulent volcanic flows (braided volcanic channels are locally observed on the Earth' moon and Mercury), the local presence of depositional bars and fans discards a volcanic origin for most martian outflow channels.

The largest martian outflow channels (Ares Vallis, Kasei Vallis, Maja Vallis, etc.) are observed at the east of the Tharsis volcanoes and usually show a depression at their head characterized by chaotic terrains (a region with collapsed plateaus and hills, Fig. 9.6e). Some outflow channels emerge out of fractures (Mangala Vallis, Athabasca Vallis). These localized valley heads explain that the term *outflow* has been used for these landforms. While still debated, these flows may have formed from the sudden release of large amount of water from the subsurface, due to ice melting and permafrost disruption. The geographic connection with volcanic landforms of the same epochs is frequent suggesting a direct relationship between outflow channels and magmatic activity. The origin of martian outflow channels is nevertheless not unique. Several landforms have been shown to be due to lake

overflows, impact craters or subglacial activity as well. Most of the observed outflow channels are Late Hesperian or Amazonian (<3.4 Gyr) and, thus, formed in periods during which Mars was no more affected by valley networks, explaining their good preservation.

9.3.3 Fluvial and Lacustrine Deposits

Transport and sedimentation of particles by river streams is an important process of sedimentation on terrestrial continents. Fluvial transport works in a relatively similar way as eolian transport, except that the fluid helps the entrainment of particles better, allowing pebbles >1 cm to be transported over >10s of kilometers. Large particles remain in the main channel forming coarse deposits while clay- and silt-size particles travel in suspension and accumulate further downward at the outlet or in adjacent flood plains. Ripples and dunes with cross-bedding features can form as well. These bedforms are somewhat similar to eolian bedforms and there is a whole terrestrial literature on the distinction between eolian and fluvial facies that can be helpful for in situ planetary observations. One easy evidence in favor of fluvial deposition is the presence of cobbles/pebbles (Fig. 9.7a) that are unlikely in the case of eolian deposits that are restricted to sand size and granule particles (<2 mm). Catastrophic floods with huge discharge rates are enable to form dunes (also named mega-ripples) made only of pebbles, as observed from glacial surges in the Scablands. Similar megaripples have been observed locally in the floor of Martian outflow channels.

Fluvial sediments accumulation is often related to plate tectonics on Earth because subsiding basins provides space to accumulate thick series of sediments. Similar processes may have existed on Mars and Titan due to other types of tectonic activity, such as regional cooling of (cryo)magmatic plumes. On Mars, channel deposits can locally be identified from the observation of inverted channels in alluvial plains. Inverted channels are formed by the preservation of the main channel deposits, with cemented coarse material being more resistant to eolian erosion than surrounding floodplain deposits (Fig. 9.7b).

The solid particles transported by rivers are stored as subaerial or subaqueous deposits after the river outlet. Alluvial fans occur in absence of standing bodies of water although shallow lakes are possible at the feet of fans (Fig. 9.7c, in contrast to delta fans Fig. 9.7d). Most typical fans occur at the feet of mountains, especially in semi-arid climate unable to generate permanent lakes. Alluvial fans are conical in 3D and are easy identifiable on images from their conical fan shape, their regular slope and the presence of multiple diverging channels on their surface. On Mars, dozens of alluvial fans have been identified throughout the planet, especially inside impact craters, which provide a steep slope and a small catchment on rims at the foot of which fans can develop (Fig. 9.7e). A famous alluvial fan on Mars is the Peace Vallis fan inside Gale crater where conglomerate deposits were identified by the Curiosity rover (Fig. 9.7a). Most alluvial fans on Earth have a mean slope of 1–5°,

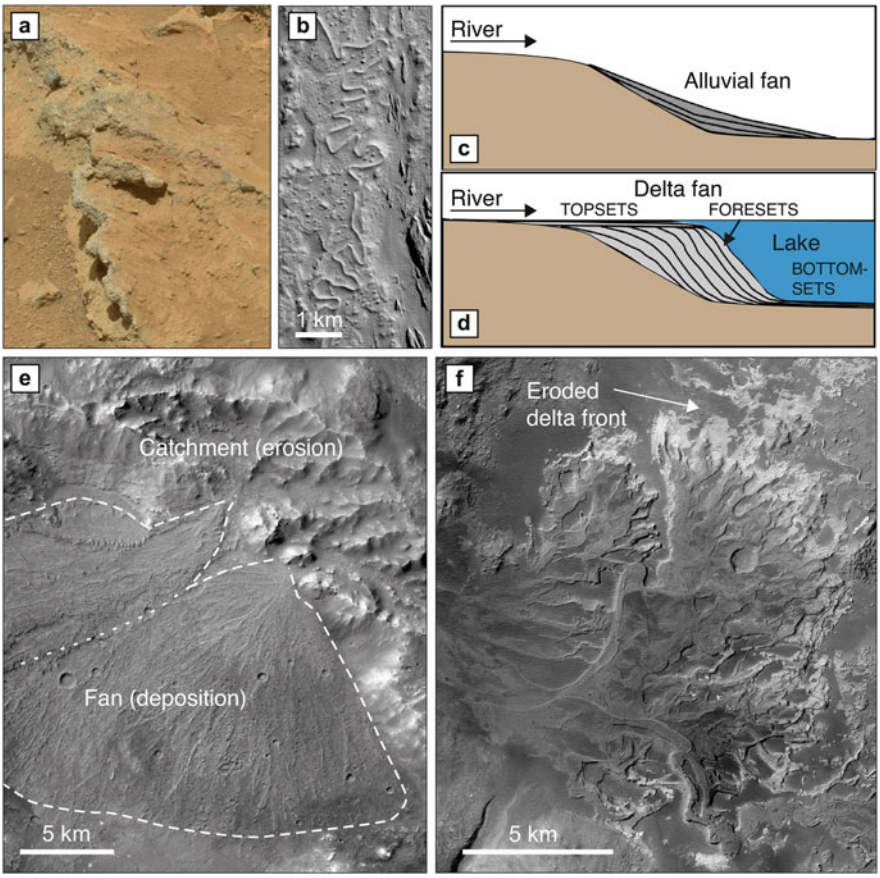

Fig. 9.7 Fluvial landforms and sediments on Mars. (**a**) Conglomerate (sediment with cemented sand grains and pebbles) observed by the Curiosity rover at the foot of an alluvial fan; (**b**) inverted channel from a meandering river in Zephyria Planum; (**c**) and (**d**) sketches of alluvial fan deposited subaerially compared to delta fan deposited below water; (**e**) two superimposed fan deposits at the base of of a crater rim, note the catchment where fluvial erosion has strongly incised the rim; (**f**) the partly eroded delta fan of Eberswalde crater with sinuous inverted channels on the delta plain and steep slopes at the front. Source: (**a**) NASA/JPL/MastCam/MSSS. (**b**) and (**e**) NASA MRO/CTX/MSSS. (**f**) ESA Mars Express HRSC

but larger fans (megafans) are also possible from the accumulation of sediments over much shallower slopes and larger catchments.

Delta fans form when rivers lead to standing bodies of water (Fig. 9.7d). Delta fans have a 3D shape distinct from alluvial fans. In most simple cases, they display a relatively flat delta plain and a steep delta front. They are often called Gilbert delta in reference to the work of Grove K. Gilbert, one of the first geomorphologists. This geometry is related to the entrainment of small-size particles by the water of the channel entering into the lake. Delta fans grow progressively keeping a relatively

similar shape. Deposits on the flat plain are usually called *topsets*, those on the delta front, *foresets*, and those on the lake bottom, *bottomsets* (Fig. 9.7d). The latter are formed by the deposition of the finest particles by suspension. This simple geometry must be taken with cautious because delta fans can build very differently depending on local context. For instance, sediments in marine context on Earth are modified by tides, sea currents and storms. The term delta, thus, includes depositional bodies from simple fan structure to large complex architecture of multiple interfingered depositional lobes.

Subaqueous deposits generate a strong interest from the planetary community because of their exobiological potential. Presumed delta fans on Mars are numerous, especially in impact craters into which a river flowed. Detailed studies of delta fans indicate local paleolakes of several 10s to 100s of meters over periods of time >10,000 years (Fig. 9.7f). Several deltas are not constrained inside depressions and could suggest larger standing bodies of water in Hellas Planitia or in the northern plains.

9.4 Mass-Wasting Processes

9.4.1 Rockfalls: Granular Behavior

The term *mass wasting* groups together all processes related to instabilities triggered by gravity. Mass wasting has an important place in the evolution of landscapes with significant topography such as mountain ranges, volcanoes or impact craters. Glacial and fluvial erosion, although driven by gravity as well, are not considered in the same category of processes, but are able to generate mass wasting indirectly by carving deep valleys with steep sides. Mass wasting includes therefore a large number of various gravity-driven slides, such as thick rockfalls with small runout (small travel distance), rotational slides with obvious slide scars, or debris flows helped by liquid water with longer runout. On Earth, the topography of the continental shelf below the sea can also be the location of submarine mass wasting, known as turbidity currents, which are a specific type of mass flows involving high velocities ($>10\,\mathrm{m\,s^{-1}}$) and very long runout distances ($>100\,\mathrm{km}$).

Friction properties control the development of landslides. The Navier–Coulomb relation links the critical shear stress at failure τ_c to the normal stress σ (∼pressure) with two parameters internal to the material, the cohesion C, and the apparent angle of internal friction φ:

$$\tau_c = C + \sigma \tan\varphi \tag{9.10}$$

The value for φ varies from 30 to 60° for typical rocky/sedimentary material. After an instability formed, landslides can be considered as a granular material with no cohesion. On a plane of slope α, the relation then simplifies into: $\tan\alpha = \tan\varphi$.

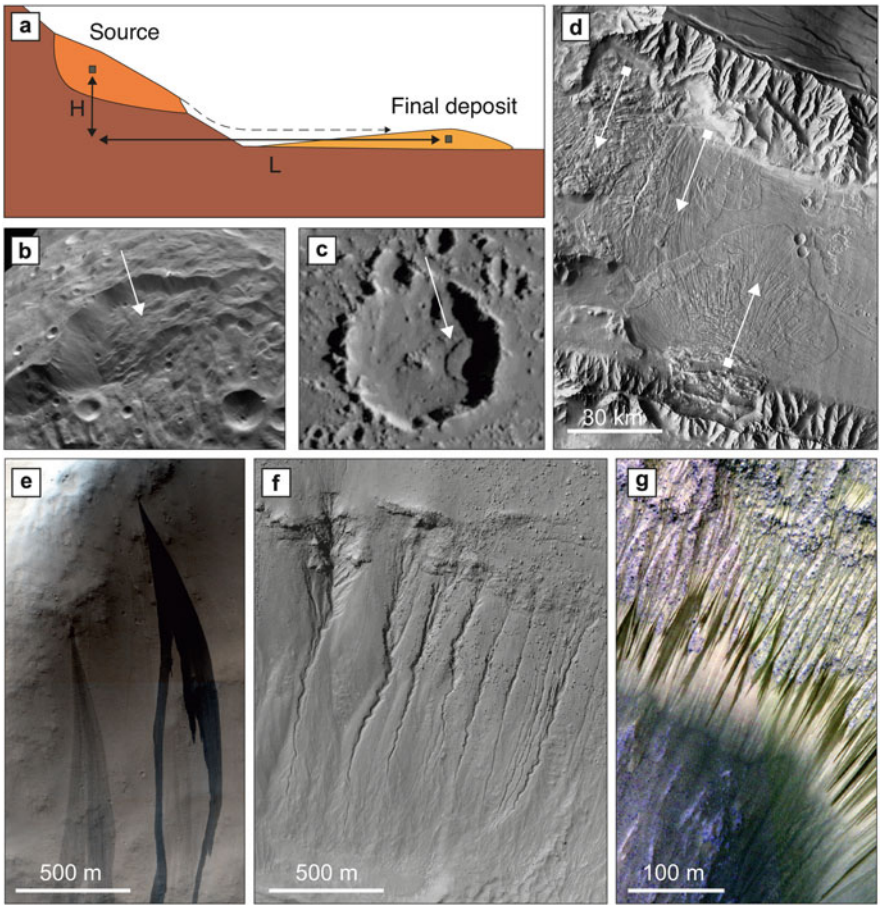

Fig. 9.8 Mass-wasting on diverse planetary bodies. (**a**) Sketch of height and runout distance estimation from large landslides center of mass; (**b**) thick landslides below a scarp on Vesta; (**c**) lobate slide inside an impact crater on Callisto. (**d**) Huge landslides in Coprates Chasma, Mars; note how the bottom one has climbed above the front of another slide; (**e**) dark slope streaks on slope inside dusty areas of Mars (NASA/MRO/HiRISE); the difference of tone is enhanced compared to reality; (**f**) recent gullies on the rim of an impact crater, note the narrow sinuous channels with small terminal deposits; (**g**) small dark streaks (recurrent slope lineae) appearing seasonally on hillslopes of equatorial and mid-latitude regions. Source: (**b**) NASA/Dawn. (**c**) NASA/JPL/Galileo Team. (**d**) NASA/THEMIS/ASU. (**e**)–(**g**) NASA MRO HiRISE Team

Thus, the slope angle can be used directly as an estimation of the friction angle of the material (referred to as apparent or effective friction angle). For large landslides on relatively flat surfaces, this relation means that $\tan\alpha = H/L$, where H and L are, respectively, the elevation difference and the distance between the center of mass of the source rock and that of the final deposit (Fig. 9.8a). Many studies have synthetized landslide properties on Earth and planets in terms of H/L-ratio,

enabling the observation of apparent friction angles. Note that, given the difficulty of finding the centers of mass, values reported usually take the overall aspect ratio of the landslides (total height and total length) rather than the center of mass, generating uncertainties on these values. While this approach remains relatively simple and valid at first order, granular materials have been intensively studied and much complex laws have been derived, involving various friction angles depending on the regime of the flow.

In the solar system, landslides are frequently observed on the rim of large impact craters on various bodies from the moon to icy satellites, independently of the presence of an atmosphere or active tectonics. These landslides are observed on planetary bodies with various gravity including asteroids and small satellites such as Vesta (Fig. 9.8b, c). The overall shape of landslides on these small bodies remains the same, showing that the process remains relatively similar. But the role (or lack of role) of the gravity on the values of friction angles is still the center of ongoing researches.

The location of the most spectacular landslides is certainly Valles Marineris, on Mars (Fig. 9.8d). The 8 km deep canyons have steep slopes that were submitted to instabilities triggered by nearby large impact craters (>10 km) or marsquakes in the Tharsis region. Landslides reach several tens of kilometers and are several tens of meters thick. Landslides scars on the scarp are up to 50 km in width. Some of the landslides have started to climb the walls of the other sides of canyons or the front of other landslides. Such observations were used to derive the velocities of these landslides calculated to be as high as 50–150 $m\,s^{-1}$. The volume generated by these landslides reaches 10^{12} m^3, higher than the largest terrestrial landslides by two orders of magnitude. The long runout distance of these landslides implies low apparent friction angles (<20°) that was initially interpreted as a consequence of internal fluids (as is frequently the case on Earth for low apparent friction angle), but more recent comparisons with numerical and experimental simulations show that no fluids are needed. This interpretation is in agreement with their chronology, because most of them formed during the 3 Gyr of the relatively cold and dry Amazonian period. Nevertheless, it remains debated if their large mobility and low apparent friction is only due to larger involved volumes or to specific fluidization processes, such as acoustic fluidization (due to vibrations) or velocity-weakening (appearing at high velocities).

Mars and the Earth's moon also display a large variety of smaller slides identified thanks to the high-resolution imagery. Among those, slope streaks are an interesting category of slides, defined by elongated tongues with smooth texture and a tone usually darker than the surroundings (bright streaks also exist) (Fig. 9.8e). These streaks were debated as potentially related to volatiles, but their shape, texture and their localization strictly linked to dust-rich areas of the planet suggest a formation by dry slides, somewhat analogous to snow avalanches. Nevertheless, the slopes of these slides are usually low, below current friction angles (<20°), suggesting that some specific processes linked to the dry fine-grained material may exist.

9.4.2 Debris Flows: Viscous Behavior

The presence of volatiles has a critical role on the triggering and the behavior of landslides. Many terrestrial landslides are triggered by heavy rains or sudden snowmelt. Water (or potentially any other fluid) present in the porosity of rocks has a direct effect on increasing the pore pressure and decreasing the effective pressure σ in relation (9.10), therefore decreasing the critical shear stress below its dry value. If water is in sufficient proportion, the mixing of solid particles and water generates a fluid with distinct properties. When water is in large amount (ca. >50% in volume), the flow is concentrated and controlled by turbulence. When water is in low proportion (ca. <10% in volume), it has minor effects on the properties of the landslides which behave as granular flows. When water proportion is between ~10–50%, the flow has particular viscous properties. These flows are usually called debris flows, although this term is somewhat used for more or less concentrated floods. *Debris flows* are usually channelized with lateral levees and a terminal deposit along or at the base of hillslopes. *Mudflow* is the term used when the concentration of fines is predominant. In the case of mudflows, the amount of water can be really low (5–10%). *Lahar* is the term used in a volcanic contact when fine ash is mixed with water. Pyroclastic flows, formed by lava eruption are not considered here but are able to display final geometry somewhat similar to debris flows.

Debris flows properties can be compared at first order to volcanic lavas, except for their thermal properties. They can be considered as one-phase fluids (i.e., debris and water form a viscous slurry) in contrast to concentrated flows, which behave as two-phase fluids (i.e., solid particles moving into water flows). The consequence is that debris flows are mostly laminar whereas concentrated flows remain turbulent. Debris flows behavior can often be approached by a simple law called *Bingham law*. A Bingham fluid is a Newtonian fluid (as defined in relation (9.3) with a threshold, or yield strength K:

$$\tau = K + \mu \frac{\mathrm{d}\gamma}{\mathrm{d}t}, \quad \text{for } \tau > K \tag{9.11}$$

$$\tau = 0, \quad \text{for } \tau < K \tag{9.12}$$

The constant K defines the strength at which the fluid starts to flow. The shear stress τ is usually related to a slope (such as $\tau = \varrho g H \sin\alpha$, with the slope α and H as the thickness of the flow), explaining that these flows are triggered on steep slopes and can stop on gentle slopes. Natural flows have more complex behavior with non-linear variations of strain rate with increasing stress (non-Newtonian fluids), and they can vary in behavior along the flow. The viscosity of debris flows is somewhat comparable to that of effusive lava flows, in the order of magnitude of 1 Pa s for muddy slurries to 1×10^5 Pa s for coarse mixing of rocky blocks in presence of a low water fraction. At the difference of most massive landslides, debris flows are much slower, in the range of 1 to 10 m s^{-1}. The yield strength K also controls the thickness of the material. It controls the formation of lateral levees that are related to the fact

that the lateral sections of the flow have a thickness below that necessary to reach the yield strength. For materials with high yield strengths, lateral levees can reach several meters thick. Materials with low yield strengths, however, may not display levees at all.

Recent gullies are small channels formed on slopes of craters or hills on Mars (Fig. 9.8f). They display characteristics similar to debris flow features on Earth: a single channel with frequent lateral levees, terminal deposits downhill or along the slope. The local observation of levees and of braided streams suggests varying amount of volatiles. Although initially interpreted as being due to subsurface water releases, it has been thoroughly demonstrated that gullies are formed by surface flows related to shallow volatiles in or on the ground. The association of gullies with ice-rich features such as polygonal cracks and small moraines suggests a formation by melting of snow/ice deposits. The presence of currently forming gullies questions nevertheless the role of current CO_2 defrosting in their formation, even if the currently forming flows are smaller than those observed all around the mid-latitudes of Mars. Recurrent slope lineae (Fig. 9.8g) consist of smaller slides formed at steep slopes resembling, at first order, dark slope streaks (Fig. 9.8e). However, in contrast to dark slope streaks, they appear seasonally in springtime suggesting a link with volatiles, such as water released from the ground or deliquescence of salts. The presence of water remains debated given the lack of features typical of water-bearing flows (channels, sinuosity, etc.). All gravity-driven flows use the same pathways, i.e. downhill, and Earth displays locations where dry slides, avalanches and debris flows can co-exist. Discriminating dry granular flows from CO_2-driven flows or water-bearing flows on another planet requires cautious observations and good constraints on physical conditions.

9.5 Ice-Related Processes and Landforms

9.5.1 Glacial Landforms

Although ices in the solar system exist for various volatiles (CO_2, SO_2, H_2O, CH_4), this section focuses mainly on landforms related to water ice. Indeed, water ice glacial landforms are by far more abundant and well-identified. Water ice displays solid-state deformation by viscous creep, which is an important constraint for erosion.

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = B\,(\tau)^n \tag{9.13}$$

where n is the non-Newtonian exponent ($n = 3$–4), and B is a temperature-dependent viscosity parameter. This relation has been calibrated by experimental studies at various temperatures and shows that the viscosity increases by almost one order of magnitude each 10 K lost. On Earth, at temperatures close to melting point, glaciers

move at 10–100 m/year, but this rate decreases to few cm/year in the coldest areas of Antarctica. Ice deformation plays a minor role at the low surface temperature of icy satellites (<150 K), while it is fundamental in their interior dynamics due to higher temperature and pressure. Hence, water ice is a major component of planetary bodies in the outer solar system, but it has a stronger role in landscape evolution when present as glaciers on rocky planets, namely on Earth and Mars.

Ice sheets are formed by the accumulation of snow in regions where the average annual temperatures are below the melting point and the accumulation of ice is higher than the ablation due to summer melting. On Mars, annual mean temperatures are below melting point everywhere. However, the low atmospheric water vapor partial pressure (0.03% of the 6 hPa of atmospheric pressure) discards the presence of ice sheets everywhere. The accumulation of ice at the surface is possible only below the frost point (~200 K in current conditions), explaining that large ice sheets are limited to polar regions. Nevertheless, tropical and mid-latitude glaciers formed in the past thanks to higher water vapor content in the atmosphere.

The topographic profile of ice caps is controlled by the viscous deformation of ice under its own weight. This property implies a profile close to a parabola for a flat surface and a lobate tongue with a thick front for glaciers on slopes. The thickness of large ice sheets (100–1000 km) is usually limited to 3–4 km, because the vertical accumulation of ice increases the basal stress that enhances the flow rate with more lateral spreading. At cold temperatures such as on Martian polar caps (<200 K), flow rates are so small that the topography can be shaped predominantly by ablation and wind circulation rather than ice flows. Polar caps on Mars, as on Earth, display internal stratifications due to the progressive accumulation of ice through long periods, typically larger than 10^5–10^6 years. Regular layering is observed for slowly moving ice sheets, as is the case on Mars. Radar data have confirmed that Martian residual caps are mainly formed by water ice mixed with dust, except a small part of the southern cap which is made of CO_2 ice.

The thermal regime of glaciers controls their ability for landscape erosion. Wet-based glaciers are defined by the presence of liquid water at the base. Such conditions are frequent on Earth, due to the combination of relatively high surface temperature and geothermal gradient, and a strong basal stress enabling water to be liquid slightly below 0 °C. Cold-based glaciers are defined by glaciers for which no water is present at the base, as is the case for current Martian glaciers (Fig. 9.9a, b). Wet-based glaciers display a strong basal slip, which increases the basal velocities and the capacity of glaciers to erode landscapes, compared to cold-based glaciers, which present only internal deformation and no slip at the base. Hence, former wet-based glaciers are easier to identify from the observation of residual landforms than dry- or cold-based glaciers. Wet-based glaciers have a strong erosional power creating U-shaped valleys when eroding into pre-existing fluvial valleys and transporting bedrock blocks at its base. The erosion of bedrock creates various patterns, including the so called *roches moutonnées* and *whalebacks*, that form over resistant obstacles in the bedrock. When eroding into sediments, the deformation of the layers can create elongated ridges known as *drumlins* that are due to the entrainment of debris into the glacial flows. Wet-based glaciers also present

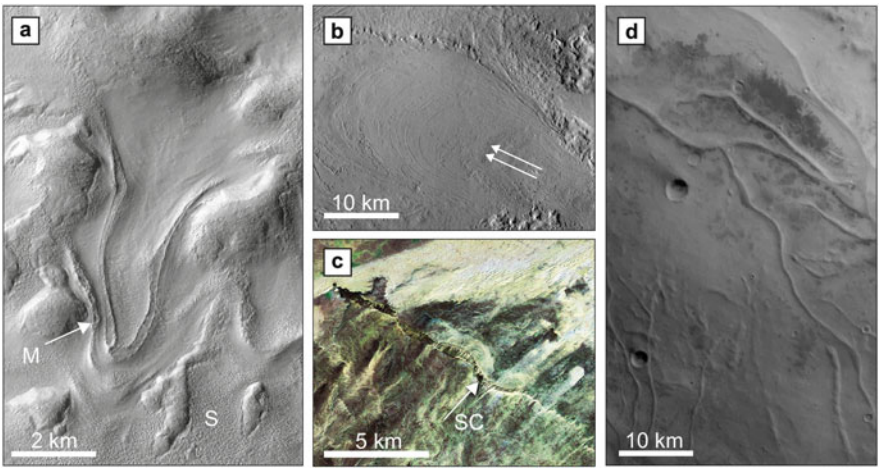

Fig. 9.9 Glacial landforms. (**a**) Glacial tongues in Greg Crater, Mars (113°E, 38°S), *M*: Moraines. *S*: terrain degraded by sublimation; (**b**) residual deposits (ablation till) from equatorial glaciers west of Olympus Mons (19°N, 220°E), arrow shows former flow direction; (**c**) Zerga mountain, a subglacial channel (SC) deposit from the Ordovician era in Mauritania; (**d**) sinuous deposits from a former ice cap (eskers) in the Hesperian period. Source: (**a**) NASA MRO/CTX/MSSS. (**b**) and (**d**) ESA Mars Express HRSC. (**c**) NASA/USGS Landsat

subglacial channels that form tunnel valleys (when eroding into bedrock) or eskers (when fluvial sediments are deposited at the surface). Tunnel valleys and eskers can be found, for instance, in Northern Canada where they formed during Pleistocene glaciations, but examples of subglacial deposits from the Ordovician glaciations (450 My) have been reported in North Africa (Fig. 9.9c). The case of eskers has been debated on Mars for decades according to images of long, anastomed ridges that mimic terrestrial eskers and that are especially observed in Argyre Planitia and around the southern polar cap (Fig. 9.9d).

Cold-based glaciers do not display any of the above mentioned landforms, but a common property of both types of glaciers is the presence of moraines (frontal, lateral or median) formed by debris pushed by ice flows. Ablation tills are poorly stratified sediments that are found as residue of the ice ablation by sublimation of cold-based glaciers. Finding glacial landforms is important to find evidence for past glaciers after they vanished. The surface of Mars presents a large variety of landforms that were once interpreted as glacial landforms, such as moraines in the northern plains or ablation tills in equatorial regions (Fig. 9.9a).

Rock glaciers are lobate tongues similar to glacial tongues but without ice visible at the surface. Rock glaciers are divided in two groups. First, rock glaciers can form by the viscous flow of permafrost or debris with ice in the porosity. Secondly, ice-cored rock glaciers or debris-covered rock glaciers are former glaciers due to snow deposition that were subsequently covered by rocky debris. The term debris-covered glaciers is specifically used on Mars for glacial tongues present in the mid-

latitudes which display a high ice content in radar data and which are covered by dust protecting them from ablation.

9.5.2 Sublimation-Driven Landforms

Sublimation is the change of state whereby a solid transform directly into a gas. On Earth, liquid water plays such an important role that sublimation is a minor process only identifiable in the driest of environments (e.g., Antarctica or on very high mountains). However, sublimation plays a strong role for both rocky and icy planetary bodies with a thin atmosphere or without any atmosphere. H_2O and CO_2 are the two main volatiles that are subject to sublimation in the conditions of planetary surfaces, but more exotic ices can locally display similar phenomena, such as SO_2 ices on Jupiter's satellite Io.

Sublimation acts differently for ground ice and for polar caps or icy satellites where ice is present directly at the surface. In the latter case, progressive sublimation can create a residual lag composed of the impurities present in the ice. Such a lag may be difficult to identify from orbital images, but it protects deeper ice from further sublimation and decrease the albedo. It has been proposed that the strong difference in albedo of Saturn's satellite Iapetus is related to a differential sublimation (Fig. 9.10a). Indeed, Iapetus' polar regions are dominated by bright terrains whereas equatorial regions are predominantly dark. This strong difference of albedo may be explained by its slow rotation that produces unusually high daytime temperatures leading to water ice sublimation in equatorial regions. Loosing water ice, these regions become darker and then also warmer, thus creating a positive feedback compared to polar regions with lighter tones. On Callisto (and in a lesser extent Ganymede), the surface presents series of pits and buttes with strong albedo variations. Buttes aligned along circles are interpreted as remnants of past impact crater rims now strongly degraded by sublimation (Fig. 9.10b).

For icy surfaces in presence of a thin atmosphere, such as on Martian polar caps, residual particles may be blown away by wind. Meter- to 10 meters-scale honeycomb-like patterns typical of ablation are observed (ablation hollows) mimicking features formed on Earth by supraglacial melting (suncups). The southern Martian cap displays a $\sim$10 m thick permanent CO_2 ice cap, which is ablated in circular or spiral shapes, usually referred to as swiss-cheese terrains (Fig. 9.10c). These geometric forms result of specific insolation conditions with low sun elevation at the pole. CO_2 has also a strong effect on landscapes during defrosting (when the seasonal CO_2 frost sublimes into the atmosphere at springtime). Defrosting creates spectacular landforms including dark spots, dark streaks and spiders—a kind of geyser related to heating and pressure build-up beneath translucent ice. The effect is especially spectacular on dark sand dunes where the difference of albedo is huge and enhances the process (Fig. 9.10d). Most of these features are ephemeral and they disappear after CO_2 ice is fully defrosted, but the long-term erosional effect of CO_2 defrosting on Martian landscape evolution is not fully understood yet.

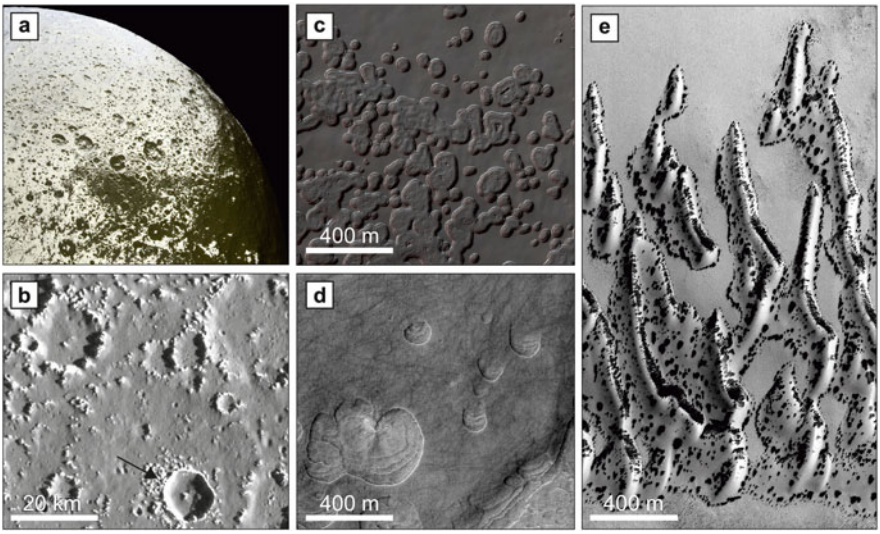

Fig. 9.10 Ice-related surface features. (**a**) Iapetus seen from Cassini; transition from bright to dark terrains; (**b**) Callisto showing terrains degraded by differential sublimation; the arrow indicates buttes of a degraded lobate ejecta; (**c**) *swiss cheese terrains* in the CO_2-covered south polar cap of Mars (86°S, 273°E); (**d**) scalloped terrains formed by sublimation of an ice-rich deposit in Utopia Planitia, Mars (46°N, 90°E); (**e**) defrosting of dark dunes at spring; bright terrains are covered by CO_2 frost (60.2°S, 7.9°E). Source: (**a**) NASA/JPL/Cassini Team. (**b**) NASA/JPL/Galileo Team. (**c**)–(**e**) NASA MRO HiRISE Team

Sublimation of ground ice is controlled by diffusion through the pores of the soil, which is linked to solid grain sizes, pores connectivity and tortuosity. Sublimation-related landforms include more or less regular patterns of networks of pits and knobs, coalescent troughs, usually at the 10–100-m scale. These landforms are sometimes referred to as cryokarst (in contrast to thermokarst usually used for volume loss due to melting). Thermal contraction cracks or glacier cracks can create fractures in the ground, which generate a regular network of troughs by differential sublimation rates. Those polygons related to sublimation are often named sublimation polygons. Strata with varying ice/dust proportions or porosities are also able to create heterogeneities in the sublimation rates, creating more or less regular patterns.

Interpreting landforms as the result of sublimation of ice from the ground is of strong interest for detecting buried ice not visible from ground penetrating radar resolution. On Mars, these landforms are ubiquitous poleward of 30° of latitude in both hemispheres. Some of the most spectacular landforms related to sublimation of water ice have been named scalloped terrains (for their resemblance to mollusk shells) (Fig. 9.10e). They may reflect the role of sublimation during a higher obliquity regime in a region of high ground ice content.

9.5.3 Periglacial Landforms

Periglacial landforms are related to cold climate independently of the presence of massive ice sheets. The permafrost is defined as permanently frozen ground, irrespective of the presence of water ice, but periglacial processes are closely connected to the presence of ice in the permafrost and the effect of diurnal or seasonal temperature variations.

Thermal contraction is one of the most developed processes involving ground ice without requiring seasonal melting. Indeed, water ice has thermal expansion one or two magnitudes larger than usual rocks. Recurrent variations in temperatures create a strong thermal stress in the ground that develops polygonal networks of cracks. Cracks often join at right angles after an evolutionary sequence in which the first set of cracks formed more or less randomly, before a second set of cracks join the first set of cracks at 90° (random orthogonal patterns). Cracks can also develop in oriented orthogonal directions (orthogonal networks) when local stress gives the starting orientation of cracks. Hexagonal polygons with 120° intersections are less frequent and are related to cracks developed at a series of points. These thermal contraction polygons attest of strong variations in temperature with presence of water ice in the ground. While cracks do not require local melting of water ice, the presence of liquid water in summer leads to the presence of ice wedges when water fills the cracks and freezes again in winter. Sand wedges can form when wind-blown sand fills cracks instead of melt water.

Polygons formed by thermal contraction cracking are typically 10–100 m in diameter on Earth, but reach 500 m in diameter on Mars (Fig. 9.11a). It has been shown that these polygons are especially abundant in high-latitude regions (>50–60°) where near-surface water ice (<1 m) has been detected by Neutron Spectroscopy. Most of these cracks are organized as random orthogonal polygons. The Phoenix lander has locally proven the presence of ice in a site with abundant thermal cracks. No contraction cracks have been observed outside Earth and Mars, although such a process may exist on any ice-rich planetary body submitted to recurrent temperature variations, such as on comets or Pluto.

The *active layer* is defined as the uppermost layer of the permafrost, which is subject to annual freezing and thawing (usually around a meter thick on Earth). Processes involving liquid water start well below 0 °C due to the presence of unfrozen water. Indeed, moisture exists down to −10 °C/−20 °C (depending on solid grain size distribution) in the form of thin films at soil grain boundaries. These films create a migration of liquid water downward during summer and upward during winter, resulting in segregation of ice on given layers. The presence of landforms related to seasonal liquid water is a strong interest for understanding the climate variations (Fig. 9.11b–f).

Sorted polygons or non-sorted polygons (also named hummocks) are circular or hexagonal patterned grounds well distinct from thermal contraction cracks. They are directly related to the migration of liquid water and formation of ice-rich lenses and the presence of stones in the ground. This process can be schematized by convecting

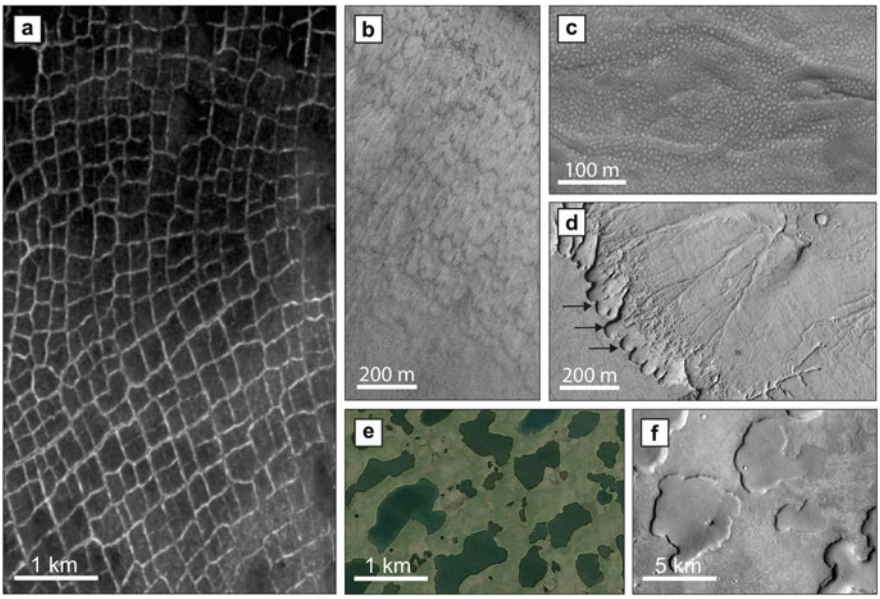

Fig. 9.11 Periglacial landforms. (**a**) Mars Observer Camera image of polygonal terrains formed by thermal contraction cracks; (**b**) Solifluction lobes on a steep slope; (**c**) Hummocky terrains covering former channels as possible indicators of freeze-thaw cycles; (**d**) A ∼10 m high scarp eroded by retrogressive thaw slumps (semi-circular features shown by *arrows*) at Cerberus Fossae, note the small channels emerging from the slumps confirming the presence of liquid water; (**e**) thermokarst lakes in the Tuktoyaktuk peninsula, Canada; (**f**) possible thermokarstic depressions in Ares Vallis signing former ice melting. Source: (**a**) NASA/MGS/MSSS. (**b**)–(**d**), (**f**) NASA MRO HiRISE Team. (**e**) NASA/USGS Landsat

cells pushing up fine grained in the center of each cell and segregating stones along their edge (for sorted polygons). These landforms are usually at the meter-scale although 20 m large patterns have been reported on Earth. On hillslopes steeper than 3–4°, these patterns transform into stripes of regular widths. They are locally associated to another periglacial hillslope process named *solifluction*, in which the active layer flows and forms meter-thick lobes.

Unfrozen water and ice segregation create buried ice lenses than can locally reach huge size (>100 m). When complete melting occurs, ice bodies in the ground melt and create >10 m large depressions, usually named *alasses*. Retrogressive slumps due to permafrost thaw can also be observed at the edge of scarps into ice-rich permafrost. Pingos are circular hills with diameters of 10–100 s of meters, 1–10 s of meters in elevation and are due to convergent migration of unfrozen water. All landforms due to segregation ice melting are grouped into the term of thermokarst because they mimic terrestrial karstic terrains due to chemical weathering of carbonates. On Titan, polar lakes were interpreted as thermokarst linked to the ablation of icy bodies, but other processes involving dissolution were also proposed.

The presence of thermokarst, sorted polygons, solifluction lobes and pingos on Mars has been debated for decades. Only the highest resolution images are able to display convincing examples of meter-scale patterned grounds. Solifluction lobes and hummocky patterns are present in the ice-rich ground of the high latitudes (Fig. 9.11b, c), possibly recording recent episodes of thaw or the presence of unfrozen water. Several examples of recent retrogressive thaw slumps (Fig. 9.11d) and sorted polygons were reported in equatorial regions (in the Cerberus Fossae region) where they may sign some unusual ground ice processes in regions expected to be currently devoid of water ice. Thermokarst due to melting of ice bodies have been observed in the northern plains and on the floor of outflow channels. In the latter case, thermokarstic depressions are linked with small channels suggesting they were filled by liquid water as observed on Earth (Fig. 9.11e, f). The role of unfrozen water in current high latitude soils on Mars is still debated after the finding by the Phoenix Lander of small amount (<1%) of carbonates and perchlorates that may be related to local chemical alteration. The identification and distribution of all periglacial landforms on a planet like Mars is fundamental for understanding the link between climate variations and water ice distribution.

9.6 Chemical Processes

9.6.1 Weathering

Weathering includes the common action of physical and chemical weathering forming soils. Mechanical weathering is the action of frost, wind or water that breaks rocks into finer grains, whereas chemical weathering transforms primary minerals into secondary minerals. The term *supergene alteration* (by opposition to hypogene alteration, at depth) is sometimes used to avoid confusion with mechanical weathering. Weathering itself has a minor effect on landscapes compared to mechanical erosion by glaciers or rivers, but it progressively transforms hard silicate grains into softer clay minerals, which in turn increases erosion rates.

The term *regolith* is used when mechanical weathering is largely predominant over chemical weathering. The word regolith comes from the word reg, meaning stone desert in the language of Sahara nomads. It has been used extensively on the moon where the regolith is produced by the prolonged action of meteorite impacts of all sizes at the surface. The *megaregolith* has been defined as the part of the upper crust gardened by impact craters over a thickness of typically 1–3 km. With the exception of planetary bodies with thick atmospheres (Venus, Earth, Titan), all planetary surfaces have a regolith due to impact craters gardening. But only few of them have chemical processes leading to the development of soils.

Pedogenesis is the usual term corresponding to the formation of soil as a consequence of weathering. Earth displays a large diversity of soils that is not detailed here. Pedogenesis is constrained by the climate, namely surface temperatures and

precipitations, the bedrock composition, and biology, on Earth. The effect of temperature on kinetics is strong: chemical alteration drops by 2 or 3 magnitudes each 10 K lost. Pedogenesis is a top-down process: The host rock—also named protolith in this context—is altered in various stages, the most altered being at the top of the sequence. The typical weathering profile starts at depth with a disaggregated host rock (or saprolite) with poor chemical modifications. Leaching of silicate minerals produce smectite clays (e.g., montmorillonite, saponite, nontronite), starting with mafic minerals such as olivine. These clay minerals still contain mobile cations such as Mg. The third stage of alteration transforms smectite clays into kaolinite clays, with a complete loss of mobile elements. Al and Fe are enriched producing various iron oxides and hydroxides (e.g., hematite, goethite) and Al-hydroxides (gibbsite) because of the lower mobility of these elements compared to the other major elements (Mg, Ca, Na, K) under near-neutral surface conditions. Laterite (also named ferricrete when reworked mechanically) is the extreme boundary of weathering on Earth, in which only Fe and Al-hydroxides survive to alteration. The thickness of weathering profiles on Earth varies from a few meters to a few tens of meters and the necessary period of time to form them can reach several million years.

Orbital spectrometers have shown that wide regions of Mars (such as Mawrth Vallis plateaus, over an area $>100\times100$ km) display continuous mineralogical sequences with widespread Fe-Mg smectite clays overlain by a $\sim$10 m thick layer dominated by Al-rich kaolinite. This sequence is found locally through the whole planet where preserved from more recent erosion, always in terrains of old ages (3.6–3.8 Gyr). The similarity with mineralogical assemblages observed in Earth for weathering profiles suggests that an intense chemical weathering occurred at that period. Ferricretes, as defined on Earth, have not been found on Mars, but locations of intense oxidation have produced high hematite layers or concretions (as on Meridiani Planum). Those localized layers may be due to other processes than weathering, such as local high oxidation states or iron-rich groundwater circulation. Whereas alteration processes on Mars become well-documented, alteration on Venus remains poorly known. The thick atmosphere of Venus made of CO_2 and acidic gases, and its high temperatures (400 °C), are conditions closer to early stages of metamorphism than usual surface conditions for alteration.

9.6.2 Authigenesis and Diagenesis

Transport of particles and dissolved ions from rivers into closed lakes can generate detrital deposits, that can be modified in place or during burial. Authigenic minerals form by alteration of silicates in situ shortly after deposition. Although many clay minerals have been identified in paleolake deposits on Mars, it is difficult to discriminate if they are authigenic—formed in situ—or if they were transported from altered upstream sources. Smectite clays found in lacustrine mudstones by the Curiosity rover on Mars are the first evidence of authigenic clays on another planet.

Diagenetic processes consist of the modification of sediments after burial by fluid circulation and enhanced pressure and temperature. This process is of great importance on Earth for the formation of oil and coal in sedimentary basins. This process is also known to generate alteration at slightly higher temperatures than surface temperatures. Hence, a significant proportion of Martian clay minerals may have formed under such conditions. Diagenetic processes in Titan's upper crust are probably important given the likely presence of a filled pore space, but remain under-constrained.

Burial also allows a strong fluid circulation through the pores of sediments or through fractures after cementation of sediments. Hydraulic fracturing is possible when the hydraulic pressure becomes higher than the strength of the sediments. Circulating fluids are charged in ions that will precipitate into fractures during water evaporation, forming diagenetic veins. Most typical minerals formed in diagenetic veins include silica, carbonates and sulfates. Calcium sulfates veins have been found on Mars by both Opportunity and the Curiosity rovers (Fig. 9.12a).

9.6.3 Chemical Sediments

Chemical sediments can form by precipitation of cations during the evaporation of ponding water. The type of chemical deposits vary strongly depending on incoming water chemistry, local weather conditions, bedrock composition, salinity, pH and redox states, etc. In addition, the chemistry of the brines changes as the evaporation progresses. This leads to various sequences usually dominated by carbonates, sulfates, and chloride salts on Earth. Sulfates observed by the Opportunity rover at Meridiani Planum, on Mars, have been explained as evaporite deposits reworked by eolian processes.

The presence of salts in soils can also create differences in aqueous conditions. For instance, iron sulfates, magnesium chlorides, and perchlorates, are known to display strong deliquescence. They absorb easily water molecules and this process can create brines that are stable down to $\sim$210 K. This opens the door to exotic chemical reactions into briny water in the cold conditions existing on Mars that are only found on Earth in a few locations such as the Atacama Desert or Antarctica. The detection of perchlorates in soils analyzed by both the Curiosity rover and the Phoenix Lander may be related to such cold climate chemical processes. Chemical processes at low temperatures are not able to explain the formation of clay minerals, but the extent and variety of minerals formed by precipitation with transient briny water in cold conditions remain unknown.

Evaporite sediments are easy to dissolve by fresh water. Karst landscapes are well-known on Earth where they form by the dissolution of chemical sediments, such as carbonates and salts. Karst landscapes are characterized by sinkholes, caves, underground drainage systems and a variety of surface patterns related to the enlargement of fractures by dissolution (Fig. 9.12b). Although this process is frequent in terrestrial conditions, the pitted texture of some sulfate-bearing layered

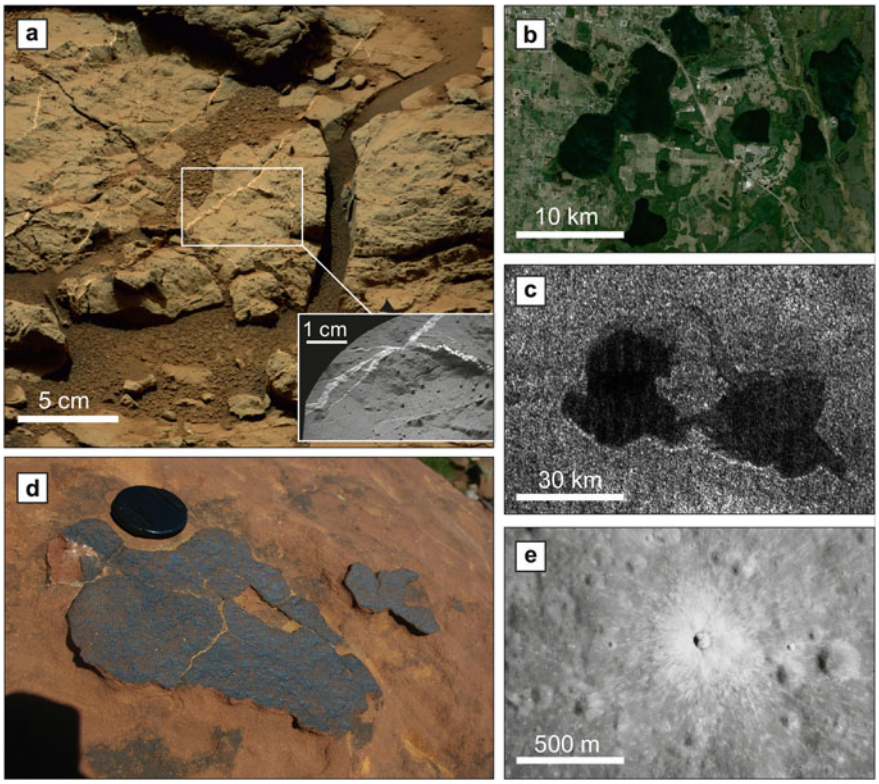

Fig. 9.12 Surface features formed by chemical processes at various scales. (**a**) Mosaic from the mudstones containing 20% of smectites, of likely authigenic origin, observed by Curiosity at Gale crater, light-toned veins are Ca-sulfate veins formed by fluid circulation after the cementation of the sediments (Mastcam image sol 126 with ChemCam/RMI insert); (**b**) lakes from sinkholes in Florida; (**c**) twin lakes at Titan polar regions suspected to be due to dissolution; (**d**) Metallic varnish in Morocco; (**e**) Small, fresh crater west of Isaev and Gagarin Craters showing bright rays over darker regolith. Source: (**a**) NASA/JPL/MSSS and NASA/LANL/IRAP/LPG/ChemCam. (**b**) NASA/USGS Landsat. (**c**) NASA/Cassini team. (**d**) N. Mangold. (**e**) NASA/JSC/Arizona State University

deposits on Mars—such as the so-called *etched terrains*—may be related to a similar process, but the involvement of water in this case in unclear. Titan's cold surface presents hundreds of lacustrine depressions with many of them nearly circular or coalescent. By analogy with terrestrial sinkholes, it has been proposed that they are related to dissolution of Titan upper crust containing complex organic compounds subject to dissolution (Fig. 9.12c).

9.6.4 Varnishes and Space Weathering

Chemical processes at the scale of individual grain size can play a strong role in changing surface properties, such as color and albedo. A coating occurs when the outer part of a grain presents a composition different from the grain itself. Coatings are important to take into account in infrared spectrometry. While coatings can correspond to any form of precipitates independent of the grain over which it forms, varnishes represent a type of coatings directly related to the bulk composition of the grain. Terrestrial varnishes are formed by $<100\,\mu$m thick films present at the surface of rocks. A variety of varnish compositions has been reported including clay minerals, carbonates, iron oxides and silica, but the most frequent varnish compositions are manganese-rich, reaching up to 50 times the proportion of Mn present in the bulk rock. Manganese is oxidized from Mn^{2+} to Mn^{4+}, which gives the black color to the common desert varnish (Fig. 9.12d). Although the name desert varnish is frequently used, similar varnishes have been observed in basically all climate conditions. Dry regions are the locations with the best preservation.

Varnish formation is still debated on Earth because of the possible role of bacteria. Indeed, the presence of Mn-oxidizing organisms has been reported frequently. The presence of such a varnish on Mars would be of strong interest for putative life forms detection. However, a pure chemical process may be possible as well. Bacteria are present almost everywhere on Earth, discarding the possibility to really test abiotic conditions. So, as a conservative assumption, Mn-rich varnishes found on Mars should be taken as an indicator of strong surface oxidation, before any further evidence of life is found.

A specific type of coating has been found on planetary bodies without atmosphere. Minerals at the surface of atmosphere-free planetary bodies are affected by micrometeorites, solar wind and cosmic-ray bombardments. These various processes are grouped under the name of space weathering—although maturation has been used as well. Space weathering has been discovered on the Moon after Apollo exploration. It was found that the surface of the regolith displays spectral properties different from the buried regolith, especially a darker albedo and subdued absorption bands. A typical example of this process is shown by the bright rays of young craters that expose unweathered material above a regolith exposed at the surface for longer times (Fig. 9.12e). These rays will progressively darken as well. The origin of this difference is a submicroscopic coating of metallic iron (Fe0), also reported as nanophase metallic iron. The formation of this Fe-coating is still debated, but the most commonly accepted scenario is a vapor condensation of Fe after sputtering by solar wind and vaporization.

Space weathering exists on most atmosphere-free bodies at various levels. Mercury, by being close to the Sun, is the most affected. It presents more craters with bright rays than the moon. Asteroids are affected as well, but with varying levels depending on their distance to the Sun, the specific asteroid class and other parameters not well identified yet. Nevertheless, the space weathering identified on Vesta by the Dawn mission has different spectral properties than the lunar spectra.

Icy bodies should theoretically display smaller space weathering given the higher distance from the sun and the lower amount of iron-rich minerals at the surface. The presence of space weathering on icy bodies studied by the Cassini probe questions their formation by other processes than solar wind. New data on asteroids and icy bodies therefore show that space weathering as learned from lunar studies does not apply the same way everywhere in the Solar System.

Take-Home Messages

Earth is the reference and serves as a natural comparison for many landforms. However, several types of landforms are not observed on Earth (CO_2 ice-rich terrains, etc.). Differences in gravity, chemistry or timescales often imply specific mechanisms. Caution must therefore be taken before applying terrestrial laws to other planets.

The presence of an atmosphere is a key aspect of the evolution of planetary surfaces by enabling transport by wind, rivers or glaciers. Without an atmosphere, surface processes are limited to processes such as vapour-solid phase exchanges or surface alteration by solar wind; endogenic and cratering processes are predominant.

Detailed topography, which is fundamental in the quantification of processes, requires stereo-imagery or laser altimetry that is only poorly available for planetary bodies other than Moon and Mars. To compensate this relative lack of information, planetary geologists should use all data available (thermal imagery, spectroscopy, etc.) that can provide useful information for geomorphology.

What is more similar to a polygonal crack formed by desiccation than a polygonal crack due to thermal contraction? How can we distinguish depressions due to dissolution (karst) from those due to melting (thermokarst)? This convergence in the shape of landforms is called *equifinality* (Chap. 2). The implication is that a given landform may not be unambiguously associated to a given process without a detailed characterization.

One key aspect of the interpretation of landforms is the understanding of the overall context, taking into account all landforms observed in the studied region and their chronological sequence of formation. Studies of planetary landforms should never consider each landform individually, but study the cortege of associated features, in a comprehensive understanding of physical and chemical conditions.

Suggested Readings

Allen, P.A.: Earth Surface Processes. Blackwell Science, Oxford (1997)

Bourke, M.C., Viles, H.A.: A Photographic Atlas of Rock Breakdown Features in Geomorphic Environments. Planetary Science Institute, Tucson (2007)

Burr, D.M., et al.: Megaflooding on Earth and Mars. Cambridge University Press, Cambridge (2009)
Carr, M.H.: Water on Mars, 229 pp. Oxford University Press, New York (1996)
Cooke, R.U., Warren, A., Goudie, A.S.: Desert Geomorphology. UCL Press Limited, London (1993)
French, H.M.: The Periglacial Environment, 2nd edn. Longman, London (1996)
Greeley, R., Iversen, J.D.: Wind as a Geological Process on Earth, Mars, Venus and Titan. Cambridge University Press, Cambridge (1985)
Greeley, R., Balme, M.R., Iversen, J.D., et al.: Martian dust devils: laboratory simulations of particle threshold. J. Geophys. Res. Planets **108**, 5041 (2003). doi:10.1029/2002JE001987
Gudipati, M.S., Castillo-Rogez, J. (eds.): The Science of Solar System Ices. Springer, New York (2013)
Knighton, D.: Fluvial Forms and Processes: A New Perspective, 383 pp. Oxford University Press, New York (1998)
Kuhn, N.: Experiments in Reduced Gravity: Sediment Settling on Mars, 167 pp. Elsevier, Amsterdam (2015)
Laity, J.: Deserts and Desert Environments. Wiley-Blackwell, Chichester (2008)
Leeder, M.: Sedimentology and Sedimentary Basins, 592 pp. Blackwell Science, Oxford (1999)
Mangold, N.: Ice sublimation as a geomorphic process: a planetary perspective. Geomorphology **126**, 1–17 (2011)
Zimbelman, J., Williams, S.: Wind streaks: geological and botanical effects on surface albedo contrast. Geomorphology **17**, 167–185 (1996)

Chapter 10
Interiors and Atmospheres

Doris Breuer and Nicola Tosi

10.1 Formation and Interior Structure of Terrestrial Bodies

10.1.1 Planet Formation

The terrestrial bodies of the solar system are the end-product of a complex process of accretion. Current theories of planetary formation predict that the birth of the Sun was accompanied by the generation of a broad, flat disk (hundreds of AU in diameter) of cold gas and dust that orbited around the newly born star owing to the conservation of the angular momentum. This *protoplanetary disk*, named solar nebula, contained all the materials from which planetary bodies formed. Although the existence of the solar nebula can only be inferred indirectly from the physical and chemical evidence left in the solar system, protoplanetary disks orbiting young stars (up to about 10 Myr old) are routinely detected via infrared- and radioastronomy, strongly supporting the hypothesis that the solar system itself evolved from similar conditions.

The time of the birth of the solar system coincides with the condensation of the very first solids from the gaseous nebula. These solids, sub-millimeter to centimeter-sized particles called calcium-aluminium-rich inclusions (or CAI) and identified in primitive meteorites, have an age of 4568 million years. The formation of solid planetary bodies, i.e. their *accretion*, is thought to consists of three main

D. Breuer (✉)
German Aerospace Center (DLR), Berlin, Germany
e-mail: doris.breuer@dlr.de

N. Tosi
German Aerospace Center (DLR), Berlin, Germany

Technische Universität Berlin, Berlin, Germany
e-mail: nicola.tosi@dlr.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_10

phases. Firstly, the collective gravity of localized aggregates of centimeter-sized dust particles leads to the formation of planetesimals, relatively small bodies that range in size from about 1 to 100 km. Once initiated, this process is expected to proceed rapidly and be completed within less than about 10^5 years. However, the occurrence in the disk of particle concentrations sufficiently large to initiate this process can be rare in space and time, with the consequence that the formation of planetesimals continues over a relatively extended period of a few million years. Secondly, within 10^5–10^6 years, planetesimals undergo a so-called oligarchic growth: large bodies grow faster than small ones because of their stronger gravity, until few tens of planetary embryos of lunar to Martian mass are formed. Finally, over a time-scale of $\sim$10–100 Myr, relatively rare collisions between planetary embryos of similar sizes, and between embryos and left-over planetesimals, give rise to a small number of terrestrial planets.

10.1.2 Interior Structure and Primary Differentiation

To the first order, the interior of terrestrial bodies, planets as well as satellites and some evolved asteroids, is characterised by a layered structure consisting of three major compositional reservoirs (Fig. 10.1 and Table 10.1). The central part is occupied by a metallic core of Fe and Ni likely alloyed with lighter elements such as S, O, and Si. The core is overlain by a rocky mantle composed of silicate minerals, with O, Mg, Si, and Fe being the most abundant elements present in the form of oxides. The crust, the outermost planetary layer, is the product of igneous processes (see Sect. 10.2.2), chemically distinct from the mantle and enriched in so-called incompatible elements. The most important of which, from the point of view of the evolution of the interior, are the long-lived radionuclides Th, U, and K. For ice-rich bodies such as the Jovian satellites Europa, Ganymede, Callisto, and the Saturnian satellite Titan, a fourth layer consisting of water/ice can exist on top of the iron-silicate layers. The mechanisms that control the formation of the above structure are strongly dependent on the thermal state of the planetary body. The differentiation of the interior into core and mantle is a simple consequence of the larger density of iron and its alloys with respect to silicates. It begins during planetary accretion and ends likely within a few million years thereafter. However, in order for metal-silicate separation to be possible, at least some degree of melting is required. Core materials, having a lower melting temperature than silicate rocks, can coexist as a liquid phase with solid silicate matrix and sink through it via grain-scale percolation or in the form of molten diapirs descending through fractures and dykes. Yet metal-silicate separation is strongly facilitated if both phases are liquid. Indeed, three main processes were likely capable to provide a sufficient amount of thermal energy to produce widespread melting during the early history of terrestrial bodies. First, the energy released by the decay of the short-lived radionuclides ^{26}Al and ^{60}Fe (with half lives of 0.7 and 2.6 Myr, respectively) caused planetesimals that accreted within the first few million years of the solar system to experience a high degree of melting.

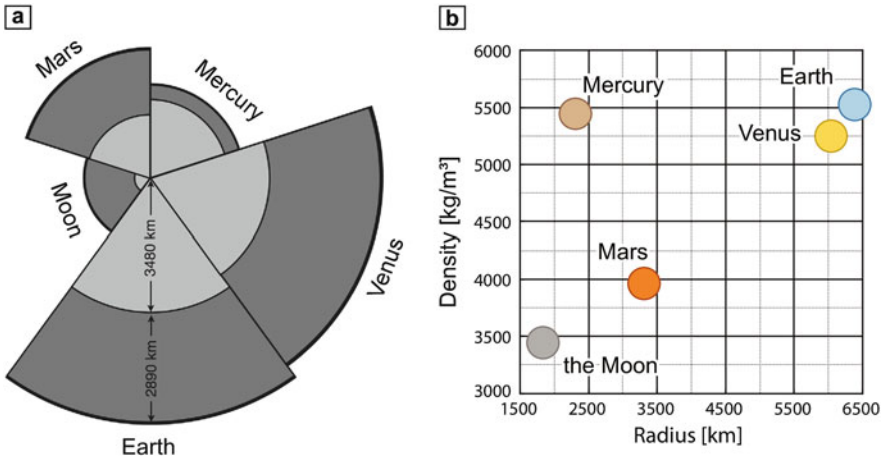

Fig. 10.1 Interior structure and mean density of the terrestrial planets and the Earth's Moon. (**a**) Basic radial structure of the interior of the terrestrial planets and of the Moon. Terrestrial bodies possess a layered structure consisting of a metallic core (*light gray layer*), a silicate mantle (*dark gray layer*), and a thin crust, chemically distinct from the mantle (*black layer* not drawn to scale). While planetary radii are known precisely, core radii (*light gray layer*) are not (see Sect. 10.1.3 and Table 10.1), apart from the case of the Earth's core, which consists of a liquid shell surrounding a solid inner core with a radius of 1220 km (not drawn).(**b**) Mean density of the terrestrial planets and of the Moon as a function of planetary radius. Note the anomalously high density of Mercury, indicative of its very large core

Table 10.1 Relevant planetary data for interior structure and evolution modelling

Planet	R (km)	R_c (km)	MOIF	Ra	$\|B\|$ (nT)	T_{surf} (C)	P_{atm} (bar)
Mercury	2440	1940–2140	0.346	Subcritical-10^5	300–400 (20)[a]	167	–
Venus	6052	3089	–	10^7–10^9	–	462	92
Earth	6371	3480	0.331	10^7–10^9	25.000–60.000 (8000)[b]	14	1.014
Moon	1737	150–400	0.392	10^5–10^8	(100)	0	–
Mars	3390	1400–1900	0.366	10^6–10^9	(12000)	−55	0.0075
Ganymede[c]	2631	700–900	0.312	Subcritical-10^6	720	−163	–

R is planetary radius, R_c the radius of the outer core, MOIF the moment of inertia factor, Ra the Rayleigh number for the mantle, B the intensity of the magnetic field either today or in () magnetic field intensity of remanently magnetized crust, T_{surf} the average surface temperature, P_{atm} the current atmospheric pressure, and X_{atm} the current atmospheric composition, with only the major constituents listed. Apart from the Earth, the size of the core is not well constrained and the values reported are only estimates based on the available data. For Venus, whose MOIF is unknown, the core radius was obtained by rescaling Earth's parameters

[a] Peak amplitudes of 20 nT observed at 15 km altitude

[b] Remanent and induced crustal field

[c] The Rayleigh number of Ganymede is given for the silicate mantle; on top of the silicate mantle is an ice layer of similar thickness $\sim$ 900 km

As a consequence, at least part of the planetesimals that contributed to the formation of embryos first and planets later were likely already differentiated, with metallic cores that contributed to build the larger cores of the bodies to which they accreted. Secondly, the late stages of accretion involved highly energetic impacts between large bodies that led to strongly increased temperatures—up to several thousand degrees—because of the conversion of the impactors kinetic energy into heat. And thirdly, the redistribution of mass associated with the very process of metal-silicate separation resulted in the release of a large amount of gravitational potential energy in the form of additional heat capable to induce temperature changes of hundreds up to few thousand degrees. Near the end of the accretion phase, terrestrial bodies were thus likely hot and chemically differentiated into a liquid metallic core and a fully—or at least largely molten—silicate mantle, part of which may well have been an actual magma ocean.

Melting and solidification in planetary mantles are not simply controlled by a single melting temperature, but by the *solidus* and *liquidus*. The first denotes the temperature at which the mineral with the lowest melting temperature among those that form the mantle mixture starts to melt. The second coincides with the melting temperature of the mineral with the highest melting point. Between solidus and liquidus, the mantle is *partially molten*. Because of the high liquidus temperature of silicate rocks, the liquid part of the mantle quickly starts to crystallize as heat is transported efficiently by convection from the interior to the surface where it is radiated to space. Since the liquidus increases with pressure more rapidly than the adiabatic temperature profile in the convecting mantle, crystallization is expected to proceed from the bottom upwards with incompatible elements that tend to be enriched in the overlying liquid phase. As a consequence, the uppermost mantle layers that crystallize near the end of the solidification period are likely to possess a different composition with respect to the underlying mantle and thus to form the first (primary) crust (see Sect. 10.2.2 for a discussion of the mechanisms that control the formation of what is called secondary crust). A relict of such a primary crust due to magma ocean crystallization can be found in the anorthositic crust on the Moon (Chaps. 8 and 11).

10.1.3 Constraints on the Interior Structure

The Earth's interior can be probed most effectively using seismic methods (Chap. 3). The inversion of the travel times of seismic waves generated by large earthquakes provides the most reliable information not only on the radially dependent structure of the interior of our planet but also, thanks to the technique of seismic tomography, on its three-dimensional density distribution. With the exception of the Moon, where a small network of seismometers installed between 1969 and 1972 by astronauts of the Apollo missions was operated until 1977, no seismic measurement has ever been conducted on the remaining terrestrial bodies of the solar system—though it should be noted that the upcoming NASA mission to Mars InSight (Interior Exploration using Seismic Investigations, Geodesy and Heat Transport), due to

launch in 2018, should land for the first time a seismometer on the red planet thus providing fundamental constraints on the structure of its interior.

In absence of seismic data, the bulk composition and radially-dependent interior structure of a planetary body can only be inferred from measurements of its mean density (Fig. 10.1b) and moment of inertia (see Table 10.1). The mean density is simply calculated from the ratio of mass and volume. The mass can be determined by measuring the revolution period of a natural or artificial satellite orbiting the body, or, when neither is available, by measuring the effects of the body's gravitational attraction on other planets or on smaller objects such as comets or asteroids passing nearby. The determination of the volume requires instead an accurate knowledge of the body's global shape, which can be obtained with various techniques: from Earth-based optical and radar observations to elevation measurements performed by radar or laser altimeters on board of orbiting spacecrafts. Knowledge of the mean density provides first order information on the composition of planetary bodies. As the density of deep interiors varies greatly with temperature and, in particular, with pressure, planetary scientists often use the concept of *uncompressed density*, which is simply the density that planetary materials would have at ambient temperature if the effects of self-compression were removed. Uncompressed densities of typical materials characterizing solid bodies of the solar system vary from $\sim$1000 kg/m^3 for ices, to $\sim$3000 kg/m^3 for silicate rocks, up to $\sim$8000 kg/m^3 for iron. As a consequence, while the relatively low mean density of the Moon, being close to that of silicates, hints at a predominantly rocky composition of the satellite's interior, Mercury, with its small size and a high mean density similar to that of the Earth (Fig. 10.1b), must be composed mostly of metal.

In order to infer the depth of the main compositional discontinuities (Fig. 10.1a), besides the mean density, a precise knowledge of the axial moment of inertia (MOI) is necessary. The MOI can be inferred from the long-wavelength gravity field under the assumption that the body is in hydrostatic equilibrium. However, this is a strict requirement that is approximately satisfied only by few bodies, including the Earth and some of the satellites of Jupiter and Saturn. For bodies that have a large obliquity, i.e. whose spin axis is tilted with respect to the orbital plane, and that also rotate fast enough to possess a significant flattening, the axial moment of inertia can be determined from the precession rate of the spin axis. The precession, the continuous change in the orientation of the rotation axis, in fact, arises as a consequence of the torque exerted by external bodies such as the Moon or the Sun on the rotationally-induced equatorial bulge. Measuring the precession rate is a complex task that requires long-term tracking of the position of landers on the planetary surface. Besides the Earth's, the precession rate of Mars is known from Doppler tracking of the Viking and Pathfinder landers and, more recently, of Spirit and Opportunity. Accurate measurements of the librations, small variations in the rotational state induced by external torques, when combined with measurements of the gravity field, also allow the axial moment of inertia to be determined as has been done for the Moon and Mercury. The MOI of Venus is presently unknown; the very slow retrograde rotation of the planet along with the small tilt of its rotation axis render a determination of the MOI from the rotational state very unlikely.

As a measure of the resistance to changes in rotation, the MOI is a good indicator of how mass is radially distributed in the interior. For a homogeneous sphere, the so-called moment of inertia factor (MOIF) about the rotation axis is

$$\mathrm{MOIF} = \frac{C}{M R^2} = 0.4, \tag{10.1}$$

where C is the polar moment of inertia, and M and R the planetary mass and radius, respectively. As shown in Table 10.1, the MOIFs of the terrestrial bodies are all smaller than 0.4, which indicates a concentration of mass toward their centers, i.e. the presence of a dense core. One exception to this rule is the Moon. With a moment of inertia factor as large as 0.392, the Moon's interior could also be explained in terms of a purely rocky composition. Yet seismic evidence suggests the existence of a small core, whose size, however, is not well constrained and can range from ∼150 to 400 km. Even in the simplest case of a two-layer model consisting of mantle and core only, the problem of inferring the interior structure suffers from an intrinsic non-uniqueness. Three unknowns, the density of the core (ρ_c), the radius of the core (R_c), and the density of the mantle (ρ_m), must be calculated from only two equations that relate them to the mean density ($\bar{\rho}$) and MOIF:

$$\bar{\rho} = \rho_m + (\rho_c - \rho_m)\left(\frac{R_c}{R}\right)^3 \tag{10.2}$$

$$\mathrm{MOIF} = \frac{2}{5}\left(\frac{\rho_m}{\bar{\rho}} + \frac{\rho_c - \rho_m}{\bar{\rho}}\left(\frac{R_c}{R}\right)^5\right), \tag{10.3}$$

where R is the (known) planetary radius. Figure 10.2 exemplifies the degeneracy of two-layer interior models for Mars and Mercury. The densities of the core and mantle as obtained from Eqs. (10.2) and (10.3) using the observed mean density and MOIF (see Table 10.1) of the two planets are plotted as a function of the relative core radius R_c/R. For a given choice of one of the three unknowns, these diagrams allow us to read out the value of the two remaining ones that satisfy the data. In particular, the model is highly sensitive to the choice of the core density. Plausible densities of core materials at the pressures of Mars' and Mercury's interiors range from ∼6500 to 8500 kg/m^3 depending on the amount of light elements alloyed with iron. Assuming, for instance, a density of 8000 kg/m^3, the model predicts mantle densities of about 3470 and 3400 kg/m^3, and relative core radii of 0.45 and 0.74, respectively for Mars and Mercury. Despite the simplicity of this two-layer model, the above inferred core radii are in good agreement with those obtained from significantly more sophisticated interior structure models based on a self-consistent treatment of the density distribution of core and mantle materials at high pressure and temperature according to appropriate equations of state. It should be also noted that since both models shown in Fig. 10.2 have MOIFs smaller than 0.4, solutions with a uniform planet density are not possible. As a consequence, as the relative core radius R_c/R approaches zero or one, the two constraints on the mean density and

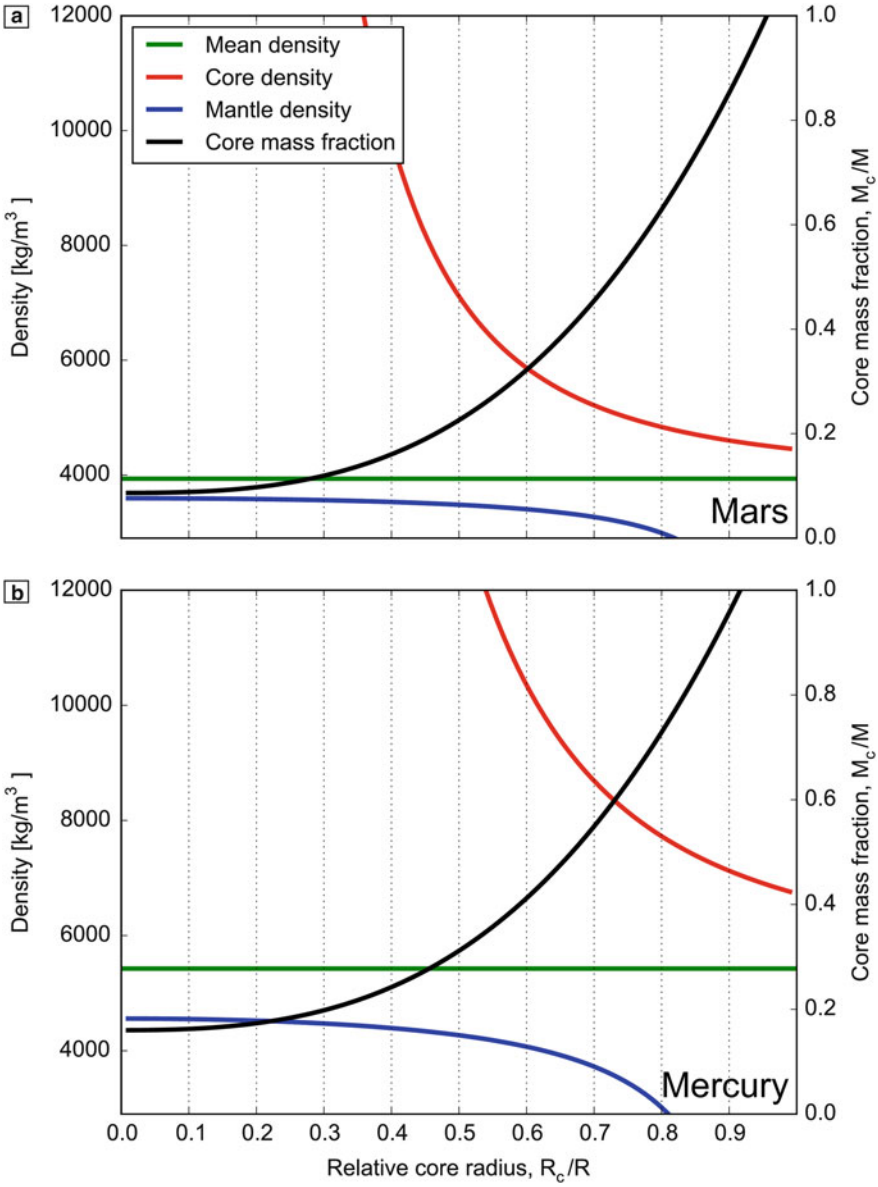

Fig. 10.2 Two-layer interior structure model of (**a**) Mars and (**b**) Mercury constrained by mean density (*green horizontal line*) and axial moment of inertia. For a given value, e.g., of the core density (*red line*), the model shows the corresponding mantle density (*blue line*), core mass fraction (*black line*), and relative core radius (*horizontal axis*) compatible with the two constraints

moment of inertia can only be satisfied simultaneously by unrealistically large core densities and unrealistically small (even negative) mantle densities, respectively.

10.2 Long-Term Evolution of the Interior

10.2.1 Convection and Rock Rheology

The completion of the large-scale solidification of the mantle sets the end of the global cataclysmic events associated with planetary accretion and with the primary differentiation of the interior. The time to reach this stage can vary considerably among different bodies but an upper limit is likely set by the giant Moon-forming impact constrained at ∼100 Myr (after the condensation of CAIs) and followed by the solidification of the terrestrial and lunar mantles that terminated within few ten million years afterwards. The subsequent thermal evolution of the interior is governed by the balance between the heat released by the decay of long-lived radiogenic isotopes present in the mantle and crust and the heat that escapes the planet and is lost to space. In order to understand this balance it is thus necessary to characterize the processes that control the way heat is transported from the interior to the surface.

The principal mechanism through which planetary bodies tend to lose their primordial and radiogenic heat is *convection*. As opposed to *conduction* that transports heat at microscopic level by transferring kinetic energy between neighboring molecules, convection involves heat transport by bulk material motion. Convection in fluids is based on the simple concept that an increase in temperature of a fluid parcel produces a reduction in density due to volumetric thermal expansion. If the fluid is subject to a gravitational field, buoyancy forces arise that tend to lift the heated parcel, with the opposite argument that clearly applies in the presence of a temperature decrease.

Let us consider a homogeneous plane layer of thickness D whose upper and lower surfaces are maintained at constant temperatures T_0 and T_1 with $T_1 > T_0$. Owing to the higher temperature, fluid near the hotter lower boundary is relatively light and will tend to rise, while fluid near the colder upper boundary will tend to sink. In order for the fluid to actually become unstable, i.e. for motion to set in, the fluid's viscous resistance must be overcome. It can be shown that only one non-dimensional parameter is sufficient to characterize the stability of a fluid heated from below and cooled from above in a gravitational field. This parameter is the *Rayleigh number*:

$$Ra = \frac{\rho^2 c_p g (T_1 - T_0) \alpha D^3}{k \eta}, \tag{10.4}$$

where T_1, T_0 and D have been defined above, ρ is the fluid's density, c_p the isobaric heat capacity, g the gravitational acceleration, α the coefficient of thermal expansion, k the thermal conductivity, and η the dynamic viscosity. The Rayleigh number describes the ratio between buoyancy forces due to thermal expansion and contraction that tend to promote convection and viscous forces that tend to suppress

it. If Ra is greater than a certain critical threshold Ra_{crit}, perturbations applied to a fluid initially at rest will be amplified and grow to a motion of finite amplitude. If instead $Ra < Ra_{\mathrm{crit}}$, perturbations will decay and energy transport will take place by conduction. The exact value of Ra_{crit} depends on the geometry of the convecting layer, on whether this is heated from below, from within or both, and on boundary conditions. Nevertheless for the plane layer configuration described above as well as for more complex ones, Ra_{crit} is about 10^3.

Although rocky mantles are by all means solid in that they propagate elastic waves on short timescales, on geological timescales of the order of millions of years, the thermally activated migration of crystalline defects allows mantle rocks to deform plastically like extremely viscous liquids in a way similar to how glaciers flow on timescales of years. The viscosity is a fundamental material parameter that measures the resistance of a substance to deform in response to applied stresses. A typical value of the viscosity of dry rocks at conditions of the present-day Earth's upper mantle is $\sim 10^{21}$ Pa s, an enormous number when compared to the viscosity of ordinary fluids—at ambient conditions, water has a viscosity of $\sim 10^{-3}$ Pa s, honey of ~ 10 Pa s and ice of $\sim 10^{13}$ Pa s. Setting in Eq. (10.4) parameter values appropriate for the Earth ($\rho = 3300\,\mathrm{kg/m^3}$, $c_p = 1200\,\mathrm{J/(kg\,K)}$, $g = 9.8\,\mathrm{m/s^2}$, $T_1 - T_0 = 3000\,\mathrm{K}$, $\alpha = 3 \times 10^{-5}$ 1/K, $D = 2890\,\mathrm{km}$, $k = 3\,\mathrm{W/(m\,K)}$) along with $\eta = 10^{21}$ Pa s for the viscosity, we obtain $Ra \sim 10^8$, much greater than the critical value for the onset of convection. In this sense, the Earth's mantle is not an exception and all terrestrial planets, possibly apart from Mercury at present-day, have a mantle characterized by a largely supercritical Rayleigh number, which is indicative of a convecting interior. In Table 10.1, we listed ranges of possible Rayleigh numbers for the terrestrial planets, the Moon and Ganymede. On the one hand, differences in the Rayleigh number among different bodies are primarily due to the thickness of their mantles, which enters the Rayleigh number with the third power [Eq. (10.4)]. On the other hand, the ranges reported for each body are a consequence of relatively large uncertainties that affect the estimate of the viscosity.

Theory and laboratory experiments indicate that the viscosity of rocks is a complex, non-linear function that can depend on several parameters such as mineral composition, stress, grain size, water and melt content. In first approximation, however, the viscosity is largely controlled by an Arrhenius-type exponential dependence on temperature T and pressure P:

$$\eta = \eta_{\mathrm{ref}} \exp\left(\frac{E^* + P\,V^*}{R\,T} - \frac{E^* + P_{\mathrm{ref}} V^*}{R\,T_{\mathrm{ref}}}\right), \tag{10.5}$$

where η_{ref} is a reference viscosity attained at reference temperature and pressure T_{ref} and P_{ref}, R is the well known gas constant, and E^* and V^* are activation energy and volume, respectively. The first accounts for the energy necessary to form vacancies and for the energy barrier that atoms must overcome to migrate into a vacancy site; the second for the fact that pressure tends to render the two processes of vacancy formation and migration more difficult.

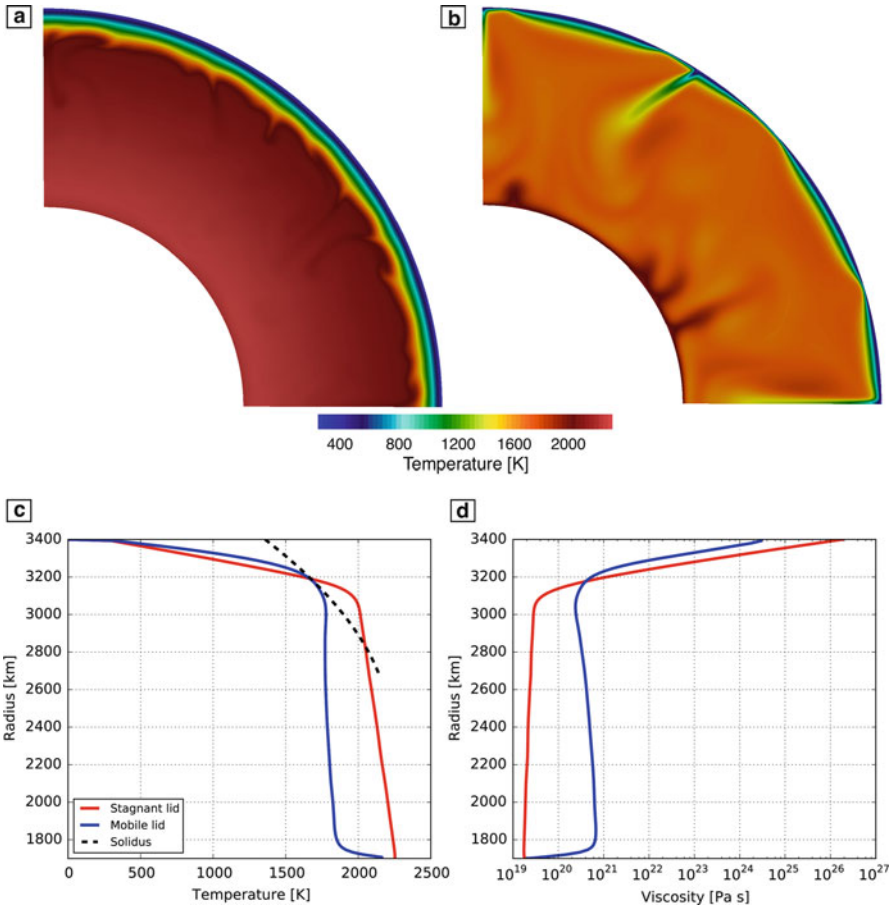

Fig. 10.3 Convection in a stagnant-lid and mobile-lid regime. (**a**) Snapshot of the temperature field from a numerical simulation of convection in the mantle of Mars assuming a stagnant lid regime, as appropriate for the present-day Mars and (**b**) a hypothetical mobile-lid regime as if Mars surface were characterised by Earth-like plate tectonics; (**c**) corresponding profiles of temperature and (**d**) viscosity. The *black line* in panel (**c**) indicates the solidus temperature

A fluid layer in which convection is fully developed is characterised by a well defined thermal profile. Temperature gradients are concentrated in two relatively thin *thermal boundary layers* across which heat is transported by conduction. The bulk of the convecting layer instead, in which temperature is efficiently advected by fluid motion, is essentially isentropic and the average temperature increases with depth because of adiabatic compression. This configuration is portrayed in Fig. 10.3 that shows the temperature field obtained from two numerical simulations of convection in the mantle of a Mars-like planet along with the corresponding laterally averaged profiles of temperature and viscosity. Because of its exponential dependence, the viscosity of cold, near surface layers increases dramatically. As

a consequence, purely viscous media whose rheology is simply controlled by temperature and pressure as in Eq. (10.5), tend to develop a thick, immobile lid—termed *stagnant lid*—which does not participate in the convection (Fig. 10.3a). This situation is the most common among terrestrial bodies of the solar system. Mercury, Mars, and the Moon possess a stagnant lid today and likely did so throughout most or even all of their history. The Earth is the only planetary body presently in a *mobile lid* regime with a surface split into tectonic plates that are an active part of the mantle convection system. Because of the relatively young age of its surface, Venus is generally believed to have experienced an episodic regime in which short events of mobilization have been interspersed between long phases of quiescence associated with a stagnant lid, which characterises the present-day condition of the planet. The ultimate reasons for which a planet preferentially develops a stagnant rather than a mobile lid are not yet fully understood. Nevertheless, the implications of the mode of surface tectonics on the dynamics and thermal evolution of the interior are multiple and profound (Sect. 10.2.2). They are a consequence of the simple fact that in a stagnant-lid planet the cold surface and uppermost layers do not take part in the convection of the mantle and thus act as a thermal insulator that tends to keep the interior relatively warm (Fig. 10.3a). In a mobile lid planet, in contrast, these layers do participate in convection and provide an efficient way to cool the interior (Fig. 10.3b).

10.2.2 Thermal and Magmatic Evolution

As briefly described in Sect. 10.1.2, following accretion, primary differentiation and mantle solidification, planetary interiors are expected to be hot, with a mantle temperature near the *solidus* of silicates and a liquid core likely superheated (i.e. with a higher temperature) with respect to the overlying mantle because of the accumulation at depth of the heat released through the loss of gravitational potential energy due to the very process of core formation. In order to describe the thermal evolution of planetary interiors, it is necessary to quantify how the loss of such primordial heat, which is largely controlled by mantle convection, and the heat production due to the decay of long-lived radiogenic elements present in the mantle evolve over ∼4.5 billion years, the age of the solar system. In first approximation, the thermal evolution of the mantle and core can be described in terms of two simple equations governing the energy balance of the two reservoirs:

$$M_m c_m \frac{dT_m}{dt} = M_m H_m + A_{\mathrm{cmb}} q_{\mathrm{cmb}} - A_s q_s \tag{10.6}$$

$$M_c c_c \frac{dT_c}{dt} = -A_{\mathrm{cmb}} q_{\mathrm{cmb}}, \tag{10.7}$$

where the subscripts m and c refer to the mantle and core, respectively, M is the mass, c the specific heat, T the temperature, t the time, and $H = H_0 \exp(-\lambda t)$

the rate of radiogenic heat production per unit mass, which starts from the initial value H_0 at the beginning of the evolution ($t = 0$) and decreases exponentially with time according to a decay constant λ that depends on the half lives of Th (1.4×10^{10} yr), U (4.47×10^9 yr for ^{238}U and 7.04×10^8 yr for ^{235}U), and K (1.25×10^9 yr). The surface area of the planet and of the core mantle boundary are denoted by A_s and A_{cmb}, respectively, while q_s and q_{cmb} indicate the average heat fluxes across them. Equation (10.6) simply states that the time rate of change of thermal energy of the mantle is balanced by the difference between the rate of heat production due to internal heat sources ($M_m H_m$) and to the heat from the core ($A_{cmb} q_{cmb}$), and the heat that is lost at the surface ($-A_s q_s$). Note that Eq. (10.6) assumes that the mantle can be characterized by a single temperature T_m and that heat sources are uniformly distributed. The energy conservation for the core [Eq. (10.7)] is somewhat simpler in that no internal heat production is considered and the evolution of the core temperature T_c is governed by the heat flow that escapes the core ($-A_{cmb} q_{cmb}$) and enters the mantle with the opposite sign.

In order to accurately determine the heat fluxes q_s and q_{cmb}, fully dynamical convection models such as those shown in Fig. 10.3 are generally necessary. Such models, especially when ran in three-dimensional geometry, can be extremely expensive from a computational viewpoint. However, instead of resolving the flow responsible for convective heat transport, appropriate scaling laws can also be employed to parametrize it. These scaling laws, that are derived from numerical experiments in which the full dynamical and rheological complexities of the mantle are taken into account, predict a power-law dependence of the heat flux on the Rayleigh number with $q \propto Ra^{\beta}$, where the exponent β is about $1/3$, though it can slightly vary depending on the heating mode of the convective system and on whether a stagnant or mobile lid regime is considered.

In Fig. 10.4 we show the evolution of the mantle temperature (Fig. 10.4a) and viscosity (Fig. 10.4b) for a Mars-like planet obtained using a simple model based on the parametrized approach described above. The figure illustrates five models in which we varied the initial mantle temperature, the reference viscosity, and the tectonic mode, either stagnant lid, as appropriate for Mars, or mobile lid, as if Mars possessed plate tectonics. As typical for stagnant lid bodies, the thermal evolution is characterized by an initial heating phase during which internal heat production is not compensated by convective heat loss (solid lines in Fig. 10.4). However, as time evolves, internal heating decreases. Furthermore, as the temperature initially rises, the mantle viscosity diminishes [see Eq. (10.5)]. A lower viscosity implies a higher convective vigor that makes the mantle cool more efficiently. As a consequence, after $\sim$1.5–2 Gyr, convective cooling takes over and the mantle actually starts to cool. Upon mantle cooling, on the other hand, the viscosity tends to increase, reducing convective vigor and rendering heat transfer less efficient. This feedback mechanism caused by the temperature dependence of the viscosity acts as a *thermostat*: relatively small temperature changes can produce large variations in the heat flux so that, after a sufficiently long time, the temperature is buffered at nearly constant values: near the end of the evolution, mantle temperatures tend to converge despite their relatively large initial differences (compare the three solid

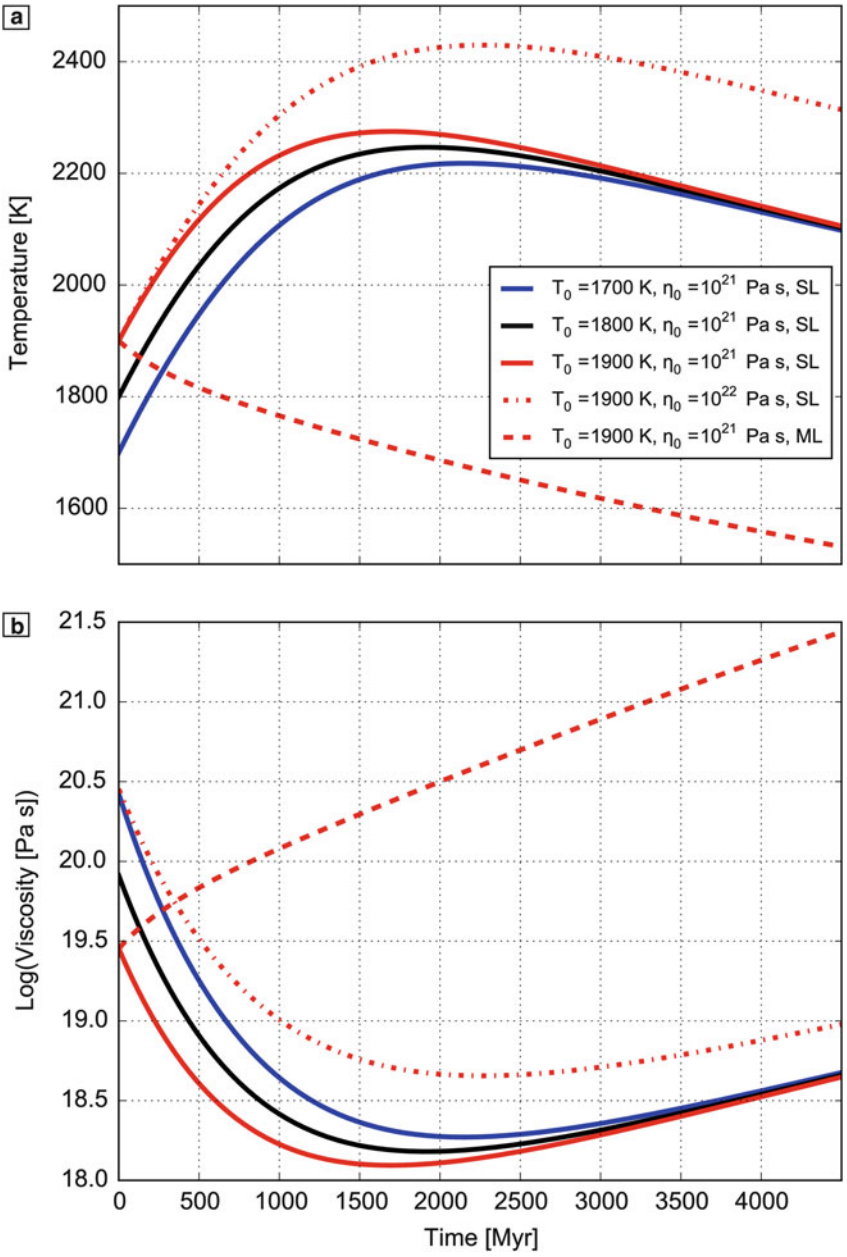

Fig. 10.4 (**a**) Mantle temperatures and (**b**) viscosity as a function of time for a simple model of parametrized thermal evolution; *solid lines* were obtained assuming a stagnant lid (SL) regime, a reference viscosity $\eta_{\text{ref}} = 10^{21}$ Pa s and initial temperatures $T_m^0 = 1700$, 1800, and 1900 K. The *dashed-dotted line* refers to a simulation with $T_m^0 = 1900$ K but a higher reference viscosity, $\eta_{\text{ref}} = 10^{22}$ Pa s. The *red dashed line* was obtained assuming $T_m^0 = 1900$ K, $\eta_{\text{ref}} = 10^{21}$ Pa s but a mobile lid (ML) regime (as if the planet had plate tectonics)

lines in Fig. 10.4). If, however, the reference viscosity of the mantle is too large, the above described thermostat effect is no longer efficient: the temperature rises more rapidly and for a longer time before heat loss can surpass heat production, with the consequence that the present-day mantle temperature turns out to be more than 200° higher than in the previous cases (dashed-dotted line in Fig. 10.4). The figure also shows how important can be the effects on the thermal evolution due to different tectonic modes. In contrast to the stagnant-lid regime, in the mobile lid regime, the cold lithosphere that is continuously recycled into the mantle cools the interior in an extremely efficient way; the initial heating phase disappears and the mantle, which cools continuously, reaches at the end of the evolution a temperature as much as 600 K lower than in the stagnant lid case (compare solid and dashed red lines in Fig. 10.4).

As the mantle heats up or when hot upwelling material rises to relatively shallow depths, the temperature can locally raise above the solidus leading to the formation of *partial melt* (see also Fig. 10.3). At relatively low pressures (less than about 10 GPa), silicate melts are less dense than solids and can buoyantly percolate via porous flow or migrate upwards forming dykes and fractures over significantly shorter timescales (up to about 10^4 years) than those that characterize subsolidus convection ($\sim 10^6$ years). Subsequent cooling and solidification of the extracted magma, either at depth or following eruption at the surface, lead to the formation of the *secondary crust*. The composition of the crust, which clearly depends from the composition of the mantle, is also representative of the degree of melting experienced by the parent material and hence it hints at the thermal state of the mantle from which the melt originated. Its thickness and time of emplacement also contain important information on the thermal history of the interior. As discussed in Sect. 10.1.2, incompatible elements are preferentially partitioned in the liquid phase. This is the case of heat producing elements, which tend to be ultimately enriched in the crust and thus depleted in the mantle. As a consequence, crust production can promote the cooling of the mantle from which internal heat sources are extracted. Volatiles, such as water and carbon dioxide, are also incompatible. Magmas enriched in volatiles can eventually feed volcanoes whose eruptions release them in the atmosphere with important implications for planetary climates (see Fig. 10.5b and Sect. 10.4).

In Fig. 10.5, we show the outcome of a rather sophisticated parametrized thermal evolution model for Mars in which not only the simple energy balances [Eqs. (10.6) and (10.7)] of mantle and core have been considered, but also the effects of melting, crustal production and partitioning of incompatible elements, radioactive heat sources as well as volatiles, have been taken into account. In Fig. 10.5a, we plotted the evolution of the mean crustal thickness, of the thickness of the stagnant lid and of the region of the mantle where the temperature lies above the solidus and hence partial melting is generated (see also Fig. 10.3). For the same simulation, Fig. 10.5b shows the evolution of the atmospheric partial pressures of H_2O and CO_2 that have been extracted from the mantle and delivered to the surface. The initial heating phase described above causes the production of a large amount of partial melt over a wide region that extends from the base of the stagnant lid

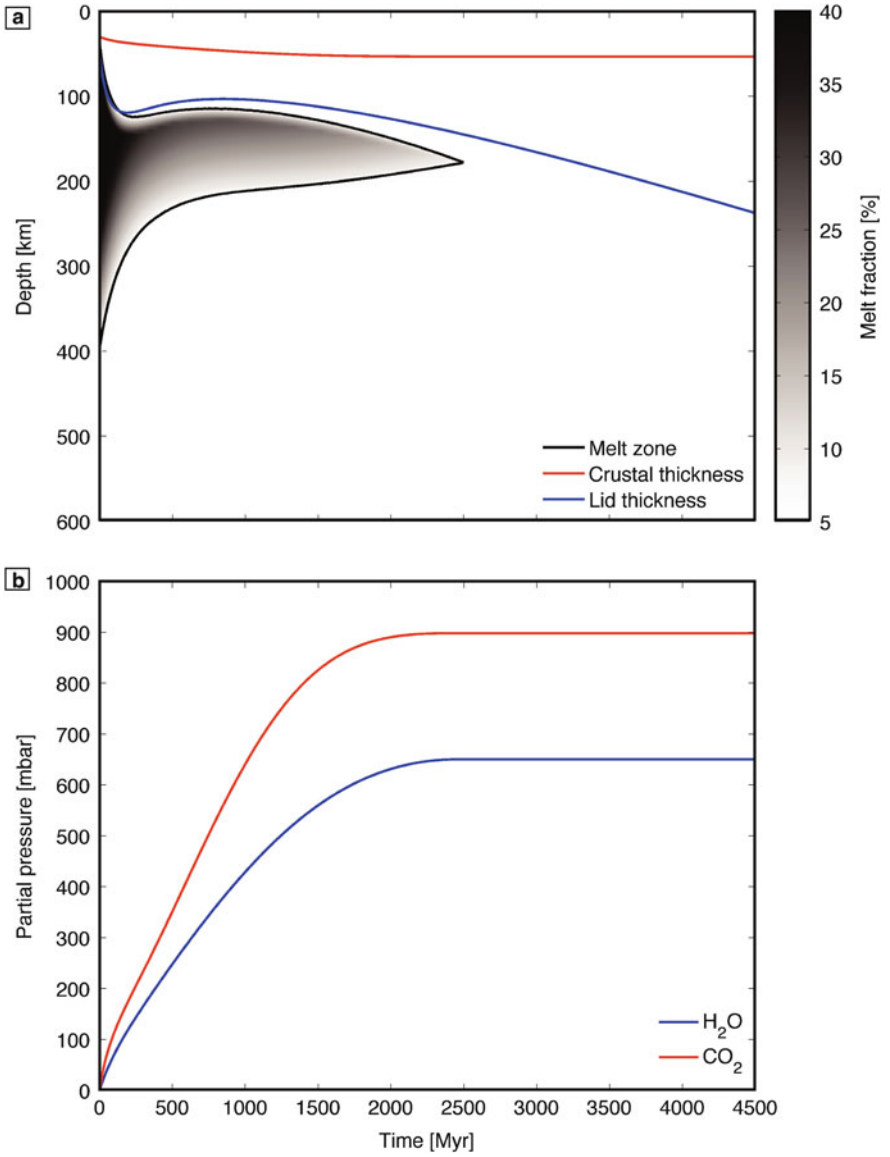

Fig. 10.5 (**a**) Time evolution of the crustal thickness (*red line*), stagnant lid thickness (*blue line*), and melt region (*shaded area*) for a thermal evolution model of Mars incorporating the effects of melting, crustal production and partitioning of incompatible elements; (**b**) corresponding atmospheric partial pressures of H_2O and CO_2 extracted from the mantle upon melting and degassed at the surface

down to a depth of up to 400 km. The bulk of the secondary crust is emplaced during the first billion year of evolution, a quite common scenario that characterizes stagnant lid bodies. The crust thickness, which starts from an initial value of 30 km assumed to correspond to the thickness of the primary layer generated during the early differentiation of the mantle (Sect. 10.1.2), reaches its maximum value of ∼50 km after ∼1.5 Gyr, in good agreement with the mean thickness of the martian crust estimated from gravity and topography measurements (Chap. 11). It is during this time that most volatiles are extracted from the mantle upon melting and melt migration and degassed into the atmosphere (Fig. 10.5b).

10.3 Magnetic Field Generation

10.3.1 Dynamo Generation

The generation of a global magnetic field requires a rotating, electrically conducting shell within a planet as well as fluid motion, i.e. convection, within that shell. In the terrestrial planets and satellites this region is believed to be the iron-rich fluid core. One way to drive convective flow is by thermal buoyancy (*thermal dynamo*), which results from a sufficiently large temperature difference between the core and the base of the surrounding rocky mantle. For convection to occur, the heat flow along the core adiabat must be exceeded. Alternatively, the growth of a solid inner core can also provide a buoyancy flux that may drive the dynamo (*chemical* or *compositional dynamo*). Compositional convection can take place when a buoyant chemical component is released during the freezing of a fluid core with a non-eutectic composition, a situation that requires the core to be composed not of only pure iron but of a mixture of iron and light alloying elements. Although there is no direct evidence for the existence of light alloying elements in the core, their presence is expected on cosmochemical grounds and has been postulated for all terrestrial planets and moons, even though their nature and concentration can considerably vary. The role of light alloying elements is twofold: they contribute to reduce the density of the core and to significantly depress its melting temperature. It must be also noted that, although the effects of a variety of different elements are conceivable, the physics of core crystallization has been studied, both experimentally and theoretically, mostly for Fe–FeS systems.

The conventional model of inner-core growth on the Fe-rich side of the eutectic—as we know it from Earth—is applicable to cases where the (positive) slope of the core liquidus is steeper than that of the adiabat, i.e. $(dT/dP)_{\mathrm{melt}} > (dT/dP)_{\mathrm{ad}}$, where T and P denote temperature and hydrostatic pressure (Fig. 10.6a). As a consequence, when the core has cooled sufficiently to reach the liquidus at the center, iron starts to crystallize and an inner core nucleus begins to form; upon further cooling, the core freezes from the inside out. As crystallization continues, the outer core becomes more and more depleted in iron and enriched in sulfur until the eutectic

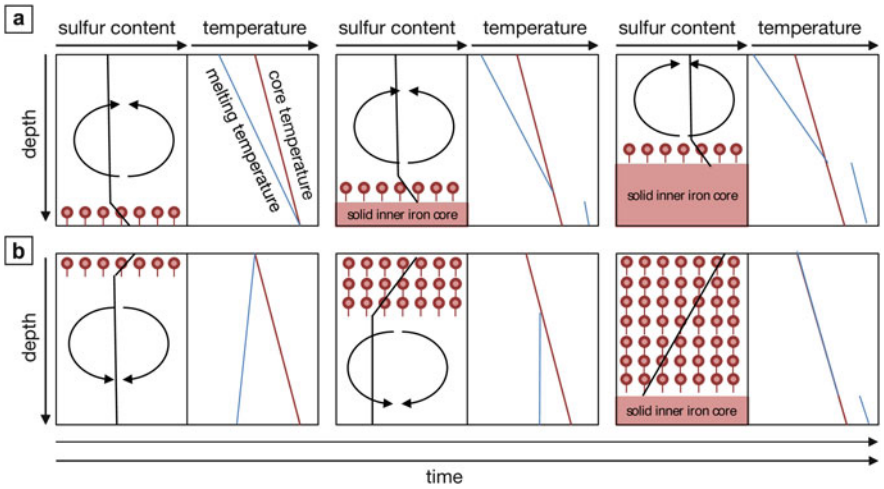

Fig. 10.6 Crystallization scenarios in the Fe–FeS system at three different points in time; (**a**) classical Earth-like case for which iron starts to precipitate at the core center forming a solid inner core; (**b**) iron snow regime for which iron crystals start to crystallize at the CMB, sink and remelt at greater depth. *Red dots* indicate solid iron. The *small dashes at the dots* mark the direction of the sinking particles. The *red solid lines* indicate the core temperature, the *blue line* the core melting temperature and the *black solid line* the concentration of sulfur as a function of depth. The *solid lines with arrows* indicate chemical convection zones. See the text for further explanations

composition is reached. Such composition is then maintained as the core continues to freeze. During crystallization, the concentration of sulfur is not homogeneous: a compositional boundary layer, in fact, forms right on top of the inner core, where the concentration of sulfur is above average. Fluid in this layer is lighter than in the rest of the fluid core, and compositional (or chemical) convection is thus induced that tends to homogenize the outer-core composition. Chemical convection and the accompanying generation of a magnetic field in the core will then occur if the temperature in the fluid (outer) core lies between that of the solidus and the liquidus of the core alloy. Note that inner-core growth can induce outer-core convection even if the heat flow through the CMB is lower than the critical heat flow for a thermal dynamo. Compositional convection, however, can be very different from the scenario described above in smaller planetary bodies. Experimental studies have revealed two important aspects of the Fe–FeS phase diagram: (1) at pressures below 14 GPa, the eutectic melting temperature decreases with increasing pressure; and (2) the S content at the eutectic decreases with increasing pressure up to a pressure of 60 GPa. Therefore, for a melting temperature profile with a negative slope (or a shallower melting temperature profile than that of the core temperature, i.e. $(dT/dP)_{\mathrm{melt}} < (dT/dP)_{\mathrm{ad}}$), Fe will precipitate at the CMB rather than in the center of the planet, with the consequence that it will sink in the form of *iron snow* because of its higher density with respect to the surrounding Fe–FeS fluid (Fig. 10.6b). A

snow zone, limited from below, forms at a depth where the core temperature is higher than the melting temperature and the iron-snow particles remelt. In the fluid where the snow zone forms, the sulfur concentration decreases with depth, which implies the presence of a stable chemical gradient across the snow zone. Remelting of iron particles below the snow zone results in a local enrichment of iron in the Fe–FeS melt. This creates a gravitationally unstable situation, where a heavier layer is formed on top of a lighter one. The Fe-rich layer initiates compositional convection and thus dynamo action. As cooling proceeds, the snow zone grows at the expense of the deeper, well-mixed fluid core. Finally, an inner core begins to nucleate when the snow zone encompasses the entire core. A dynamo would thus only be active during the period between the formation of the snow zone and the time when it reached the center of the core.

Regardless of whether chemical or thermal buoyancy drive core convection, the existence of a dynamo and the strength of the resulting magnetic field are strongly dependent on the rate of heat transfer rate through the mantle of the planet. On the one hand, the thermal buoyancy flux is directly related to the rate at which the mantle extracts heat from the core. On the other hand, for a compositionally-driven dynamo, the rate of crystallization (or inner core growth) and hence the rate of buoyancy release depend on the core cooling rate and therefore on the heat flow from the core. This is directly dependent on the vigour of mantle convection and on mantle cooling rate, whose evolution can be calculated from models of thermal history such as those presented in Sect. 10.2.2.

Thermal evolution models suggest that the existence of a present-day magnetic field in a terrestrial planet requires a crystallizing core. A purely thermally driven dynamo is unlikely for any present-day terrestrial body since models typically predict a slowly cooling core with a heat flux that is smaller than that conducted along the core adiabat. In fact, a thermal dynamo is generally difficult to obtain for stagnant-lid planets unless the core is initially superheated with respect to the mantle. If this is the case, e.g. following core formation (see Sect. 10.1.2), a thermal dynamo typically shuts off very early during the evolution—as inferred for Mars from the remanent magnetization of its crust—since the heat flow from the core decreases very rapidly during the first few hundred million years. Whether the core starts at all to crystallize (and by which crystallizing mechanism), and whether a compositional dynamo can be initiated depends strongly on the thermal evolution, on the pressure range covered by the core, and on its composition, both of which affect the melting temperature. For Mars, which lacks a global magnetic field at present, it is suggested that the core is still completely fluid, a scenario also hypothesized for Venus, whose slow retrograde rotation (a Venus day of $\sim$243 Earth days is almost equal to the length of its year of $\sim$224 days) is alone not sufficient to rule out a dynamo action. The mechanism of iron-snow crystallization has been proposed for Ganymede and Mercury to explain their present-day magnetic fields. In the case of Mercury, the large core and accompanying broad pressure range even suggests the simultaneous existence of multiple iron snow zones. The latter may explain its weak field strength, which indeed poses severe problem to hydromagnetic dynamo models based on the classical inner core growth mechanism.

A chemical iron-snow dynamo could also start operating on Mars in the future and has been postulated for the early dynamo of the Moon for a core sulfur content above ∼8 wt%.

10.3.2 Crustal Field

Magnetized minerals within the crust and possibly the upper mantle can also contribute to the magnetic field of a terrestrial body. A rock containing magnetic minerals carries a natural remanent magnetization (NRM) that is the vector sum of all the different possible components of magnetization acquired over its history. The presence of remanent magnetization is indicative of an earlier field generated via dynamo action. NRM can be acquired in the presence of an ambient field through the following mechanisms:

- thermoremanent magnetization (TRM). Magnetic minerals will acquire a stable TRM when cooled from a temperature initially above their Curie temperature (T_{curie}) and a partial thermoremanent magnetization (pTRM) of lower remanence, when cooled from a temperature initially smaller than T_{curie}. The slower the cooling rate the stronger the remanent magnetization will be.
- chemical remanent magnetization (CRM). When magnetic minerals form or crystallize, they may acquire a CRM.
- shock remanent magnetization (SRM). This type of magnetization can be acquired when magnetic minerals are shocked with transient peak pressures of several GPa induced by impacts.

Note that the above mechanisms may be also responsible for demagnetization if no ambient field is present. As a prominent example, the magnetic field signature of large demagnetized impact basins on Mars has been used to constrain the end of the Martian dynamo era, which likely lasted for the first 500 Myr of evolution. In a similar way, the magnetization of the crusts of Mercury and the Moon suggests the presence of an early self-generated magnetic field though, in contrast to Mars, their dynamos persisted longer: for Mercury until present, while for the Moon until 3.2–1.3 Ga before present at a level of a few μT—though, in this respect, the available data do not provide a conclusive estimate. Therefore, the level of magnetization of the crust and the time at which crustal materials formed and cooled below their Curie temperature (or the time of an impact) can be used to assess whether and when an internally self-generated magnetic field was present or not.

There is a trade-off between the concentration of magnetic carriers and the strength of the magnetic field: the lower the magnetic field, the more magnetic carriers are required to explain an observed magnetization. In particular, ferro- and ferrimagnetic minerals are the most suitable candidates to account for the stability and strength of the Martian crustal field (Fig. 10.7) and, among these, magnetite, hematite, and pyrrhotite with Curie temperatures of 850, 940, and 600 K, respectively. The surface temperature of Venus at about 740 K is relatively close to

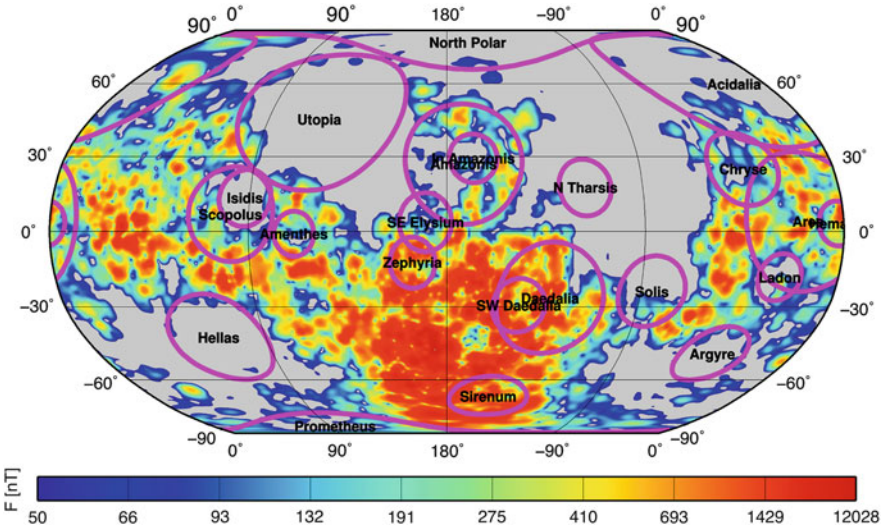

Fig. 10.7 Intensity of the Martian lithospheric magnetic field evaluated at the mean planetary radius of 3393.5 km. Major impact basins larger than 924 km in diameter are indicated by *thick purple circles* and labelled with their names. Source: courtesy of A. Morschhauser

T_{curie} of the main magnetic carriers. Therefore, the expected temperatures of Venus' crust, possibly with the exception of a thin upper crustal layer, are likely above the Curie point. Any remnant crustal magnetic field from an early period of dynamo activity should thus be weak on Venus if existent at all. For icy satellites, any possible remanently magnetized basaltic crust is located underneath the ice layer—in the case of Ganymede below an ice mantle of about 900 km in thickness. As the magnetic field strength is inversely proportional to the cube of the distance from the magnetic anomaly, the associated signal will be very difficult to detect from orbit.

In addition to the time constraints of an ambient magnetic field, the crustal magnetic field allows us to study tectonic processes such as the movement of the magnetic pole with time—the so called true polar wander, i.e. the shift of the geographical poles relative to the planetary surface. In both cases, it is necessary to determine the direction of the magnetic field with respect to the direction of the geographic coordinates. Although orbital measurements are referenced to geographic coordinates, it is highly challenging to resolve NRM at the stratigraphic bedrock scale and to correlate this to well defined geologic processes. For instance, magnetic stripes observed in the Mars crustal field that exhibit a magnetic field of alternating sign have been interpreted in terms of a possible form of seafloor spreading (Fig. 10.7). Such stripes can be revealed at an orbiter height of 400 km but it is uncertain whether the same structure would also persist at a smaller scale. A high resolution is thus crucial for the interpretation of magnetic field data. Paleomagnetic data relative to isolated rocks and boulders are most useful as they allow for the unique determination of the net magnetic moment of the samples and,

in turn, of the position of paleopoles. However, with the exception of analyses of lunar samples collected during the Apollo missions, so far, direct measurements of the magnetization of individual, isolated rocks are restricted to meteorites whose geologic context is unknown. Alternatively, the magnetic moments and paleopoles can be estimated by modeling magnetic field anomalies of known source geometry, with the latter that can be inferred from gravity data.

10.4 Planetary Atmospheres

10.4.1 Composition and Surface Temperature

In contrast to the Earth's atmosphere, which consists primarily of N_2 and O_2 and the most abundant trace gases H_2O, Ar and CO_2, the composition of the atmospheres of Mars and Venus are dominated by CO_2 ($\sim$95–97 vol.%). The next most abundant gas is nitrogen for both planets ($\sim$3%), while trace gases are Ar, CO, H_2O and O_2. On Venus, small amounts of SO_2, H_2SO_4 and HCl, and on Mars of ozone have also been detected. While the surface pressure on Mars is only $\sim$7 mbar, a dense atmosphere with a surface pressure of 93 bar exists on Venus. It has been estimated that the amount of CO_2 stored in near-surface layers on Earth in the form of carbonates, if released, would be equivalent to the amount of CO_2 in the present atmosphere of Venus. Mercury and the Moon, as well as most of the icy satellites, only have very tenuous atmospheres. One exception is Titan with a N_2-dominated atmosphere and a surface pressure of 1.5 bars.

The dense CO_2-rich atmosphere on Venus is also responsible for the efficient greenhouse effect that rises the surface temperature above the planetary equilibrium temperature by about 500 °C to an average value of 462 °C. Additional greenhouse gases are H_2O, methane and ozone. For Mars with its thin atmosphere, the greenhouse effect is much less efficient and leads to a temperature increase of less than 10 °C that translates into an average temperature of −55 °C. Such temperature along with the low atmospheric pressure do not allow fluid water to be stable on the Martian surface. In absence of a strongly insulating atmosphere such as the one of Venus, lateral variations in surface temperature can be very prominent. On Mars, the equatorial temperature is by as much as $\sim$170 °C greater than the temperature at the cold poles. On Mercury, with its tenuous atmosphere, 3:2 spin-orbit resonance, high orbital eccentricity and nearly zero obliquity, the surface experiences large temperature differences both in latitude ($\sim$160 °C) and in longitude, resulting in equatorial hot and warm regions with a temperature difference of about 90 °C.

10.4.2 Atmosphere Formation and Loss Processes

During accretion, a proto-atmosphere is formed around a planet by capture and accumulation of gases from the nebula and by early degassing of the initial volatile inventory acquired from colliding planetesimals, impacting asteroids and comets. The processes of early degassing involves dehydration and hydrous melting, impact devolatilization and catastrophic outgassing due to magma ocean solidification, all of which can potentially lead to the early formation of a thick proto-atmosphere with a pressures in excess of several hundred bars.

Depending on the distance to the Sun, on the planetary mass and atmospheric composition, a significant part or even the entire proto-atmosphere can be lost. In general, the lower the planetary mass, the closer the planet to the Sun and the less massive the gas molecules, the greater the likelihood of escape. In particular, high EUV flux of the young Sun can induce hydrodynamic blow-off of hydrogen and strong thermal escape rates of dragged heavier species such as O and C atoms. It is suggested that Mars may have lost its entire primordial atmosphere during this stage, while Venus with its higher average density did not. More distant bodies have cooler atmospheres with a range of lower velocities and less chance of escape. For instance, this is what helps the saturnian satellite, Titan—which is small compared to Earth but further away from the Sun—to retain its atmosphere.

During the subsequent long-term evolution following planet formation and possible magma ocean solidification, a secondary atmosphere can be generated by interior outgassing or by the delivery of volatiles through late impacts. As discussed in Sect. 10.2.2, upon mantle melting, volatiles such as H_2O and CO_2 are preferentially enriched in the liquid phase. Volatile-rich partial melt, having typically a lower density than the solid mantle, rises toward the surface, forms secondary crust and redistributes the volatiles. On the one hand, as the solubility of volatiles in magmas at surface pressure is relatively low, essentially all dissolved gases are released into the atmosphere if the melt reaches the planetary surface (*extrusive volcanism*). On the other hand, if the melt does not reach the surface but rather forms a pluton at depth (*intrusive volcanism*), only part of the volatiles will eventually be delivered to the atmosphere depending on the porosity of the crust and on the depth of intrusion. It should be also noted that because of the significantly higher solubility of H_2O in melts with respect to CO_2, a high surface pressure caused by a dense atmosphere tends to inhibit efficient degassing of water but not of carbon dioxide, an effect, highly relevant for Venus, that can ultimately lead to a water-poor atmosphere in spite of a possibly long-lasting volcanic activity.

The rate of volcanic outgassing can be estimated from the rate of crustal production and from the concentration of volatiles in the extracted magma as follows:

$$\frac{dM_{\mathrm{i}}^{\mathrm{atm}}}{dt} \propto \frac{dM_{\mathrm{cr}}}{dt} \gamma X_i^{\mathrm{melt}}, \tag{10.8}$$

where $M_{\mathrm{i}}^{\mathrm{atm}}$ and X_i^{melt} are respectively the outgassed mass and the concentration of the ith volatile species, M_{cr} is the amount of produced crust, i.e. of extracted magma, and γ is a factor controlling the outgassing efficiency.

The amount of volatiles in the melt is limited either by their respective solubility or by the way these are partitioned in the liquid phase—the mechanism providing the smaller value is the limiting factor. For water, the saturation concentration is relatively high at pressures where the melt is generated and the enrichment of water into the melt is controlled by partitioning during fractional melting. Therefore, the water concentration in the liquid depends on the initial concentration in the solid phase, on the partition coefficient and on the melt fraction. In contrast, the CO_2 content in the magma can be limited by its solubility in particular for a low oxygen fugacity of the mantle. For a silicate mantle under reducing conditions, as it is suggested for Mars, carbon is expected to be stable as graphite or diamond. Accordingly, the enrichment of CO_2 in the melt is lower than what is typically expected for melt on Earth which is characterised by more oxidizing mantle conditions. For stagnant lid planets, most degassing and thus most of the atmosphere formation occurs in the first few hundred to one billion years, i.e. during the time of efficient crustal formation (Fig. 10.5b) as also indicated by their old planetary surfaces. On exception might be Venus showing a much younger surface with an average age of about 700 Ma (Chap. 11).

The atmosphere is further influenced by loss processes, which are dominated by impact erosion and by non-thermal escape due to solar wind forcing in the late evolution. The efficiency of the latter process is reduced in the presence of a magnetosphere, which protects the atmosphere by deflecting the incoming solar wind. It has been suggested that the disappearance of the early magnetic field on Mars was responsible for efficient loss of an early secondary atmosphere, resulting in the present-day thin atmosphere. Furthermore, the atmosphere can interact with crustal reservoirs by CO_2 weathering and hydration processes, which occur at the surface and/or in the crust. But also the storage and release of carbon in near-surface reservoirs such as carbonate, or ices may change the atmospheric pressure and composition.

Take-Home Messages

Terrestrial planets formed within 10–100 million years after the birth of the Sun via collisions between particles, planetesimals and planetary embryos that grew larger and larger at the expense of small bodies

Terrestrial planets are differentiated bodies with an interior structure consisting of three main reservoirs: a central metallic core composed of Fe and Ni and some light alloying elements (e.g., S, O, and Si), a rocky silicate mantle composed of oxide minerals and, on top of it, a thin crust with distinct composition that forms through igneous processes

In absence of seismological observations, the relative size of the core and mantle can be inferred from remote measurements of planetary mass, radius, and axial moment of inertia, which, in turn, requires accurate measurements of the gravity field and rotational state of the planet

A fluid in a gravitational field subject to temperature perturbations undergoes thermal convection if its Rayleigh number, a characteristic ratio of buoyancy to viscous forces, is larger than a critical threshold of 10^3. Although mantle rocks have an extremely high viscosity, planetary mantles have a largely supercritical Rayleigh number and, over geological timescales, transport heat via convection

The Earth's surface is fragmented into large tectonic plates representing a cold mobile lid that participates in convection and efficiently cools the interior being continuously recycled into the mantle. On the contrary, the other terrestrial planets are characterised by a stagnant lid, a highly viscous, immobile layer, which does not take part in the convection of the mantle and retards the cooling of the interior

The thermal evolution of the interior is controlled by a balance between the heat loss due to convection, and heat production due to the decay of radioactive elements and basal heat entering the mantle from the core

Planetary magnetic fields are generated through dynamo action due to thermal and/or compositional convection taking place in the liquid core. The former requires that the mantle cools at a sufficiently fast rate, while the latter arises because of the crystallisation of the inner core that is accompanied by the release of light elements into the liquid outer core. Besides the Earth, Mercury and Ganymede are at present the only solid bodies of the solar system having a dynamo-generated magnetic field

In absence of an active magnetic field, the existence of a dynamo-generated field that operated in the past can be inferred remotely by measuring the remnant magnetisation recorded by crustal and lithospheric rocks. Evidence of ancient magnetic fields is available for Earth, Mars and the Moon

Only four Solar System bodies possess a significant atmosphere: the Earth, Venus, Mars, and Saturn's moon Titan. The main constituents are N_2 and O_2 for the Earth, CO_2 for Venus and Mars, and N_2 for Titan

The primary atmosphere of a planetary body is generated during the early evolution through the capture of gases from the nebula, impact devolatilization, and outgassing upon mantle solidification, but it can be quickly lost because of escape processes promoted by stellar activity. The secondary atmosphere is generated instead through mantle melting accompanied by the extraction of volatiles, which are ultimately released at the surface

Acknowledgements N.T. acknowledges support from the Helmholtz Association (grant VH-NG-1017).

Suggested Readings

Breuer, D., Moore, W.: Dynamics and thermal history of the terrestrial planets, the Moon, and Io. In: Schubert, G., et al. (eds.) Physics of Terrestrial Planets and Moons, vol. 10, Chap. 8, pp. 255–305. Elsevier, Amsterdam (2015). doi:10.1016/B978-0-444-53802-4.00173-1

Breuer, D., Labrosse, S., Spohn, T.: Thermal evolution and magnetic field generation in terrestrial planets and satellites. Space Sci. Rev. **152**(1–4), 449–500 (2010). doi:10.1007/s11214-009-9587-5.

Grott, M., Morschhauser, A., Breuer, D., Hauber, E.: Volcanic outgassing of CO_2 and H_2O on Mars. Earth Planet. Sci. Lett. **308**(3), 391–400 (2011). doi:10.1016/j.epsl.2011.06.014

Morschhauser, A., Lesur, V., Grott, M.: A spherical harmonic model of the lithospheric magnetic field of Mars. J. Geophys. Res. Planets **119**(6), 1162–1188 (2014). doi:10.1002/2013JE004555

Schubert, G., Turcotte, D.L., Olson, P.: Mantle Convection in the Earth and planets, 912 pp. Cambridge University Press, Cambridge (2001)

Sohl, F., Schubert, G.: 2 – Interior structure, composition, and mineralogy of the terrestrial planets. In: Schubert, G., et al. (eds.) Physics of Terrestrial Planets and Moons, vol. 10, Chap. 2, pp. 23–64. Elsevier, Amsterdam (2015). doi:10.1016/B978-0-444-53802-4.00166-4

Spohn, T., Johnson, T., Breuer, D. (eds.): Encyclopedia of the Solar System, 1336 pp., 3rd edn. Elsevier, Oxford (2014)

Stanley, S.: Magnetic field generation in planets. In: Spohn, T., et al. (eds.) Encyclopedia of the Solar System, Chap. 6, pp. 121–136. Elsevier, Amsterdam (2014). doi:10.1016/B978-0-12-415845-0.00006-2

Tosi, N., Breuer, D., Spohn, T.: Evolution of planetary interiors. In: Spohn, T., et al. (eds.) Encyclopedia of the Solar System, Chap. 9, pp. 185–208. Elsevier, Amsterdam (2014). doi:10.1016/B978-0-12-415845-0.00009-8

Part III
Integration and Geological Syntheses

Chapter 11
The Terrestrial Planets

Angelo Pio Rossi, Stephan van Gasselt, and Harald Hiesinger

11.1 Introduction

11.1.1 Comparing Terrestrial Planets

Comparative Planetology deals with the study of similarities and differences across various Solar System bodies and on the processes acting through time over them. Some of these processes, such as asteroid and comet bombardment (Chap. 7) or space weathering (Chap. 9) act on several or all of them, some only on individual (Chaps. 8 and 9) planets, moons (Chap. 12) or small bodies (Chap. 13).

For the structure of this chapter we chose to discuss the terrestrial planets and their satellites in terms of their geological evolution with time. We start with ancient times and end with modern times. Orthogonal to this time line, we describe the processes and their effects acting on each planetary body. This approach was chosen to facilitate the comparison among the studied objects by discussing the geological evolution of a planet in the context of the others. Our home planet, Earth, will serve as a reference framework and important anchor point for our comparative planetology studies because it is the best studied object for which we

A.P. Rossi (✉)
Jacobs University Bremen, Campus Ring 1, 29795 Bremen, Germany
e-mail: an.rossi@jacobs-university.de

S. van Gasselt
National Chengchi University, No. 64, Sec 2, ZhiNan Rd., Wenshan District, Taipei 11605, Taiwan
e-mail: svg@nccu.edu.tw

H. Hiesinger
Westfälische Wilhelms-Universität Münster, Wilhelm-Klemm-Str. 10, 48149 Münster, Germany
e-mail: hiesinger@uni-muenster.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_11

have knowledge gained from hundreds of years of sample analyses, mapping efforts, drilling, mining, and remote sensing, to name only a few sources of information.

Comparing the entire suite of terrestrial planets requires to individually characterize each of them and the processes acting on them through time. However, there are several limitations involved in such a reconstruction, partially due to the discontinuous and fragmentary nature of the accessible geologic records and the lack of enough well-contextualised field data. In this context, increasingly high-resolution remote-sensing coverage is important, in order to better map the nature and the stratigraphic relationship between surface units (Chaps. 2 and 5). Such information has to be augmented by in situ observations, sample return, and ideally by human exploration.

The degree to which we know the geology of individual terrestrial planets is variable: it largely depends on the amount, quality, and resolution of available remote sensing data, as well as on a possibly available ground truth. Moreover, the reconstruction of the geological history of any planet relies on time-consuming geological mapping (Chap. 4), whose pace is different for different planets. Even for some of Earth's most covered and best mapped countries, some of the base geologic maps are decades old, and refresh times are counted in several years to decades. For planets, systematic geologic mapping is affected by time lags of the same order of magnitude. On certain planets, such as Venus, even the stratigraphy is relying in some cases on workarounds involving the relative sequence of deformational events rather than the sequence emplacement of geological units.

The Moon's and Mars' geological histories have been studied for several decades and with continuously improving data coverage (Chap. 4), crossing Petabytescale. Mercury has only recently been studied with modern remote sensing data, and detailed regional and global mapping is taking place in these years. In fact, its very first global geologic map based on data from the last decade has been recently produced. Venus is globally mapped in its basic units, but its complexity, especially for older, more heavily deformed terrains requires newer higher resolution observations (Fig. 11.1).

Terrestrial planets, also known as inner or rocky planets share several characteristics: they are mainly composed of silicates, they are located relatively close to the Sun, they show evidence for past or present vigorous volcanism and tectonicsm (Chap. 8), and at least half of them hosts or hosted a stable hydrosphere at some point during their evolution (Chap. 9). These, among other conditions (Chaps. 10 and 14), contributed to their present or potential habitability. Earth's Moon is included in this chapter because it is part of the Earth–Moon system, being genetically and partially geologically linked to the Earth from very early times. The relative sizes of terrestrial planets determine to a certain extent their respective fate, with the smaller bodies, and, in the case of Mercury, also with the smaller relative mantle thickness, ended their geological activity earlier, while larger ones continued until very recently or are still active.

The current distribution of topography (hypsometry) of the terrestrial planets reflects their integrated geological history (Fig. 11.2). Although largely driven by endogenic processes (Chap. 8), surface processes and surface-atmospheric processes interaction had a strong role in shaping them at geological timescales

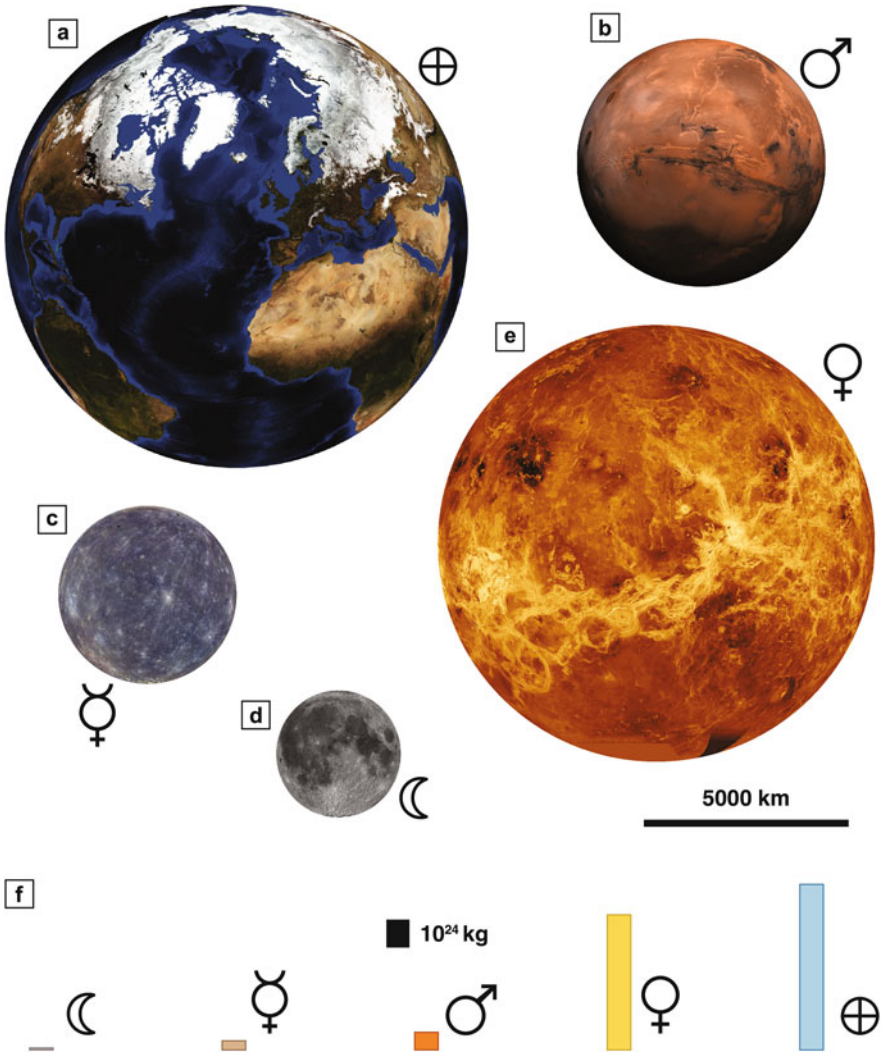

Fig. 11.1 Terrestrial planets compared, on images or renderings with negligible atmospheric masking of the surface; (**a**) Earth Blue Marble, obtained from MODIS data and bathymetry data of Earth; (**b**) Mars as seen by Viking, centered on the 5000 km long Valles Marineris canyon system; (**c**) Mercury in enhanced MESSENGER color; (**d**) The lunar nearside showing the two main different terrains, highlands and maria; (**e**) Venus, artificially colored Magellan radar backscatter image of an hemisphere. Sources: (**a**) NASA. (**b**) NASA Viking Orbiter, USGS. (**c**) NASA Messenger. (**d**) NASA/LRO/LROC. (**e**) NASA Magellan

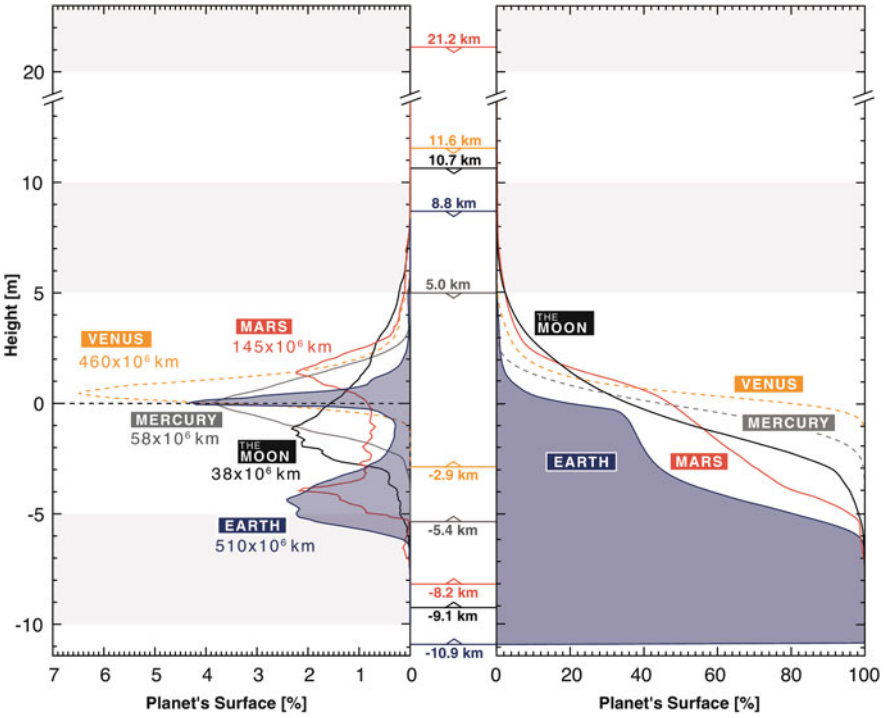

Fig. 11.2 Hypsometric curves of the terrestrial planets and the Moon. Source: Mercury: NASA/Messenger/MLA; Venus: NASA/Magellan/SAR Altimeter; Moon: NASA/LRO/LOLA, Earth: NOAA/ETOPO-1, (Amante and Eakins, 2009); Mars: NASA/MGS/MOLA

(Chaps. 9 and 10). However, the overall surface age of each terrestrial planet varies depending on its individual history (Fig. 11.3). The spatial distribution of stratigraphic units also reflect this (Fig. 11.4). When comparing the hypsometric curves of the terrestrial planets, it can easily be seen that the topographic height distribution for Earth and Mars have a distinct bimodal shape (Fig. 11.2a) while for the Moon, Mercury and Venus the distribution is unimodal and symmetric with individual maxima around the median. For Earth this characteristic shape is related to the distribution of land masses (i.e. continents) and oceans. Although is tempting to use this observation as an argument in favour of a Martian ocean, the bimodal shape is more likely related to a large impact or the result of mantle convection (Chap. 8), differentiating northern lowlands for southern highlands. In the cumulative view, the martian bimodal shape is less well pronounced when compared to Earth.

With the exception of Venus, the impact histories of the terrestrial planets are similar to a great extent, although with exact timing possible variations, in their first few hundred million years, but dramatically diverge afterwards. Such divergence

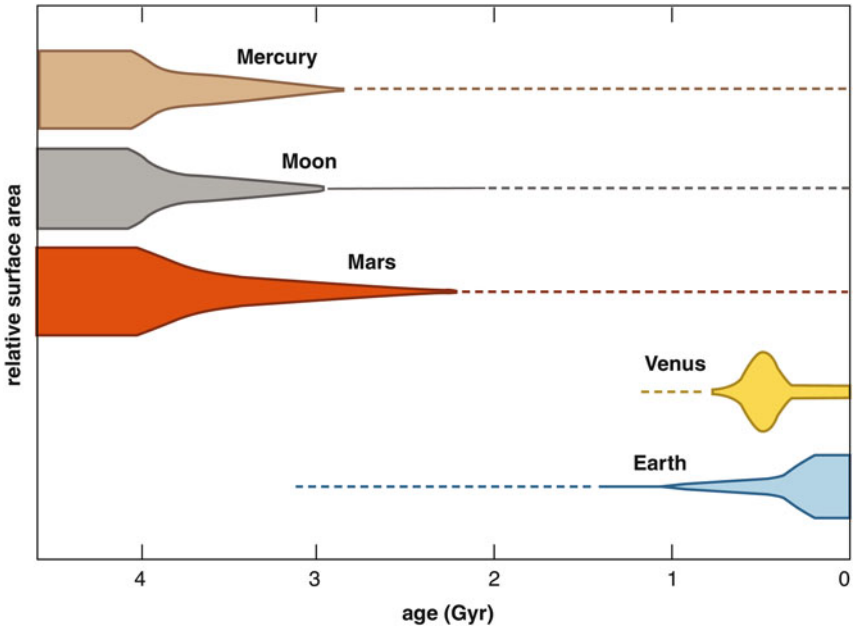

Fig. 11.3 Evolution of surface area age for all terrestrial planets and the Moon through time. Source: Redrawn from Head (1999)

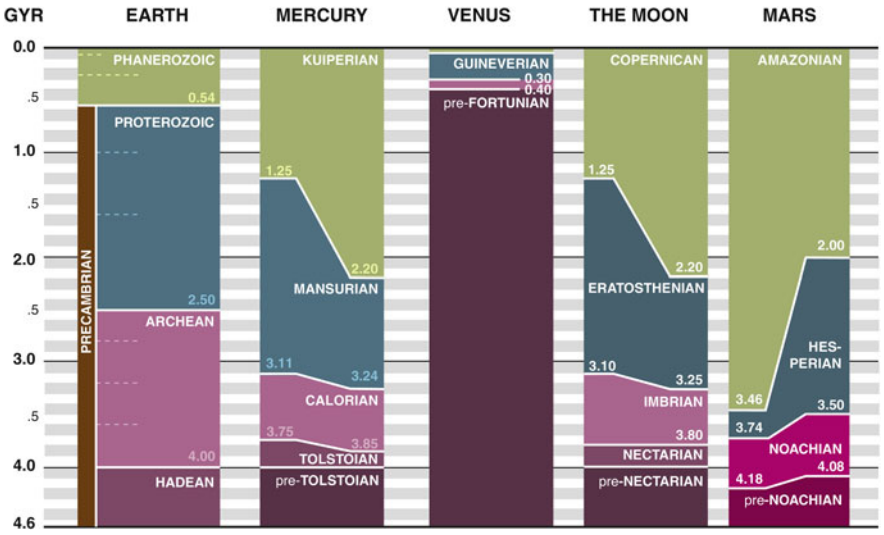

Fig. 11.4 Chronostratigraphic comparison of the Terrestrial planets. Source: Modified after van Gasselt and Neukum (2011)

is not related to the external impact cratering dynamics, but mainly to internal dynamics causing partial to total obliteration of earlier cratering record (Chap. 1).

11.1.2 Timing of Events: Cratering Histories

The stratigraphies of all terrestrial planets and the Moon have been established and calibrated, respectively, with impact crater size-frequency distribution measurements on their surfaces and radiometric dating (Chaps. 2 and 7). The latter has only been possible by returning samples from well-characterized landing sites on the Moon that allowed for the correlation of cumulative crater frequencies and the radiometric/exposure ages of the samples. This chronology curve can be adapted to other planets to date their surfaces. On Mars, a well-dated candidate crater has been proposed as source of the Shergottite-type meteorites (Chap. 6). If correct, this would indirectly provide a calibration point for the martian chronology. Another indirect absolute age determination on Mars has been achieved by NASA's Mars Science Laboratory (MSL) rover, finding ages consistent with those of the area dated with crater size-frequency on remote sensing orbital data (Fig. 11.4).

The relative preservation of surface units belonging to a certain chronostratigraphic subdivision (Fig. 11.4) depends much on the dynamics and the geological activity on each planet or moon: Extremely old terrains, formed during the earliest phases of crust formation are preserved (Fig. 11.5) on large parts of the Moon (Pre-Nectarian to Nectarian), Mercury (Pre-Tolstojan, Tolstojan) and to a significant portion of Mars' surface (Noachian).

Correlating events and evolution across the terrestrial planets is difficult and for reasons of processes specific to each planet it is not necessarily meaningful. However, impact processes in the Solar System, with their variations through space and time shaped all rocky planets in a first order similar way, particularly in the first few hundred million years (Fig. 11.5), although synchronicity might not have been exact. The impact flux (Chap. 7) had its source mostly in the asteroid belt with smaller contributions from comets (Fig. 11.6).

Other endogenic and surface processes also occurred on more than one planet, but at different times and with slightly to largely different boundary conditions. For example, early Mars is often compared to ancient Earth in terms of liquid water availability and results from the MSL rover suggest large reservoirs of water and possibly O_2in the atmosphere or in near-surface environments.

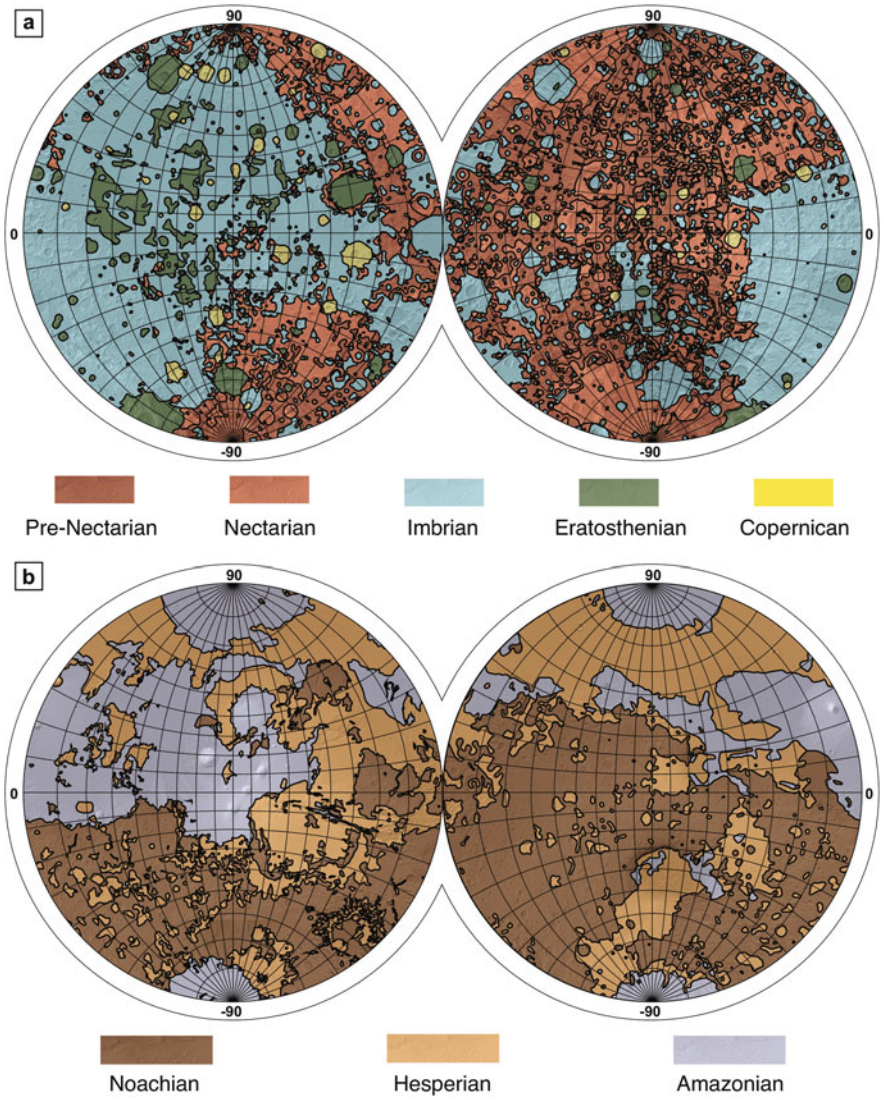

Fig. 11.5 Global surface ages, based on Fig. 11.4. (**a**) The Moon. (**b**) Mars. Sources: (**a**) After Fortezo and Hare (2013) and references therein, also quoted in Fig. 8.12. (**b**) After Tanaka et al. (2014)

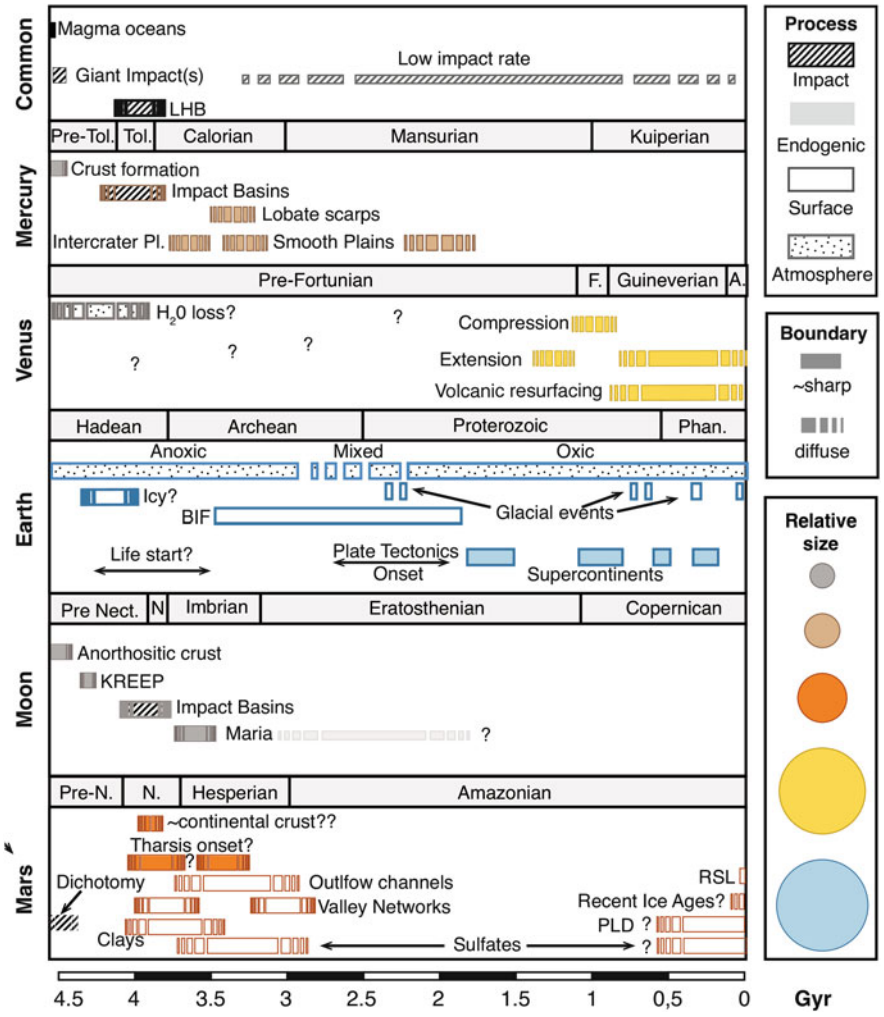

Fig. 11.6 Synoptic view of progesses acting on Terrestrial planets through time (see Chaps. 7–9): Dominating processes through time for the terrestrial planets and the Moon. See also Fig. 14.2. Source: Art and Nisbet (2012); Shearer et al. (2006); Nance et al. (2014); Basilevsky and Head (1998); Ehlmann et al. (2011); Fassett and Head (2008); Carr and Head (2010); Hoffman and Schrag (2002); Wilhelms et al. (1987); Neukum et al. (2001); Sautter et al. (2015); Head et al. (2007); Van Kranendonk et al. (2012); de Kock et al. (2009)

11.2 Early Phases

11.2.1 Formation and Magma Oceans

The very earliest times of Solar System formation are recorded only in meteorites (Chap. 6). Early solid aggregates in the Solar System were micrometersized. Collisions among those particles progressively produced larger bodies, reaching tens to hundreds of kilometers, forming so-called *planetesimals* (Chap. 1). Continuing collisions of these relatively large bodies resulted in both catastrophic disruption (Chap. 7) and planetesimal fragmentation as well as accretion into even larger, 10^3 km-sized planetary bodies, eventually leading to the formation of the terrestrial planets over time.

The initially violent phases of the formation of terrestrial planets (Chap. 1) led to global impact-generated melting involving a substantial thickness, producing so-called magma oceans. Those magma oceans affected also planetesimals in the first few million years of Solar System evolution. In the case of the terrestrial planets, it is assumed that the depth of these magma oceans that were produced by global melting were on the order of tens to hundreds of kilometers.

The subsequent cooling of a magma ocean (Fig. 11.7) led to the formation of a primary crust: its closest and most well-preserved example are the bright highlands of the Moon, composed of relatively light-toned anorthosite, as compared to darker maria basalts (Chap. 8). Individual crystals (zircons) formed during the cooling of Earth's magma ocean are as old as the Hadean (Fig. 11.4). However, most of the early geologic record is lost on Earth due to plate tectonics, the atmosphere, and life. Thus, with these factors missing, the Moon is an excellent body to study the very earliest history of the Earth–Moon system by investigating its primary crust.

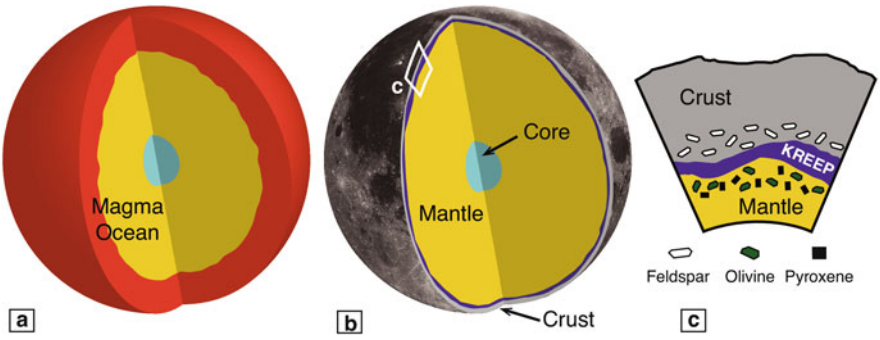

Fig. 11.7 Magma ocean of the Moon: (**a**) Initial state of lunar Magma ocean following its accretion after the giant impact; (**b**) final state of the Moon, with the original *primary* crust solidified from the magma ocean producing light-coloured, anorthositic highlands; (**c**) In the process incompatible elements are concentrated in the so-called KREEP layer, evident in the area of Oceanus Procellarum (Procellarum KREEP Terrain). Source: redrawn after Geiss and Rossi (2013)

Magma ocean conditions lasted between a few to several million years. Earth's crust existed most likely already 4.4 Gyr ago and the magma ocean, following the Moon forming impact around ~4.5 Gyr ago lasted most likely for no more than 100 Myr. The crystallisation of the lunar magma ocean, that was likely to be at least few hundred km deep, left anorthositic highlands, which are still preserved. The presence of the ancient magma ocean on the Moon is also supported by the unusual concentration of incompatible elements in certain terrains, reflecting a reservoir at depth of KREEP (K, Rare Earth Elements, Phosphorous). The relative concentration of those elements on the surface of the Moon could be mapped using Gamma Ray spectrometry. Although their concentration occurred at depth, KREEP materials were brought to the surface through later volcanism.

The inaccessibility of Venus' most ancient geological record (Figs. 11.4 and 11.6) does not allow for collecting evidence of its potential magma ocean. Mercury's very think mantle (Chap. 10) provides some constraint on the eventual size of its ancient magma ocean. Mars' magma ocean occurred early, few million years after the formation of the Solar System. The southern highlands of Mars represent the oldest portion of its preserved crust (Fig. 11.5), but only recently and only very small outcrops of anorthosites have been found on Mars, mainly in deep units uplifted by central peaks in large craters. Due to its dynamic nature, Earth lost all direct evidence of its magma ocean. However, it likely had a magma ocean at least several tens of kilometers deep.

11.2.2 Giant Impacts

The first phase of intense impact flux on the terrestrial planets is also known as the period of *Early Bombardment*. Impacts of bodies of different sizes, up to planetesimals of several hundred kilometers were common during the formation of the terrestrial planets. Large and catastrophic impact events during the first several million years are also recorded in the most ancient meteorites (Chap. 6). Such giant impacts differ from basin-forming events in scale, producing such large-scale damages that globally affected the planetary body or even disrupted it. The giant impact that likely resulted in the formation of the Moon is a good example. The giant impact hypothesis of the origin of the Moon predicts that some 50–100 Myr after the formation of Earth (4.5 Gyr), a Mars-sized body, *Theia*, impacted Earth and caused substantial damage but did not fully disrupt Earth: the resulting debris re-accreted to form Earth's Moon (Fig. 11.8).

The giant impact hyphothesis, as of today the most widely accepted for the origin of the Moon, could also be used to explain the structure of Mercury. Such an impact on Mercury could have resulted in a relatively thin mantle and a large iron core (Chap. 10). Giant impacts were large enough to be able to completely disrupt or globally affect the involved planetary bodies, with variable final fate of impactor and target depending on relative mass and encounter geometry.

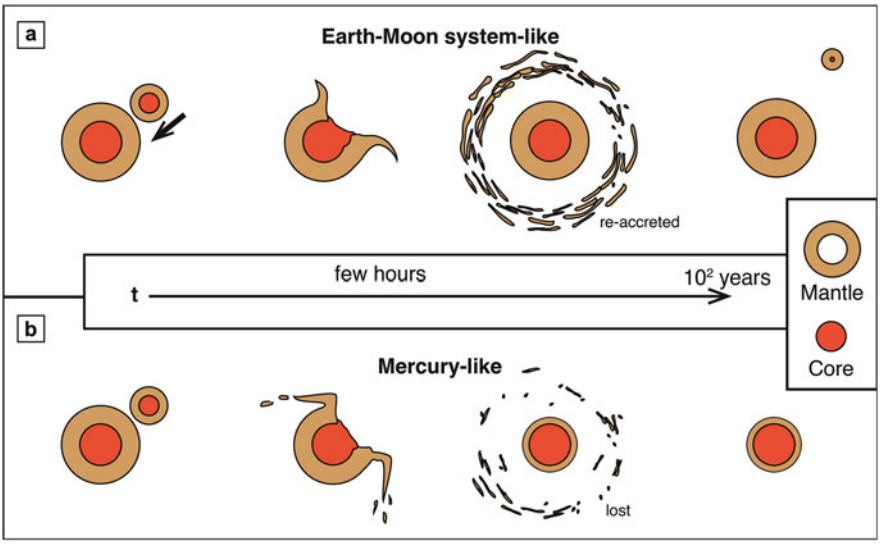

Fig. 11.8 Giant impact, roughly exemplifying the possible role of an impact of a large planetesimal on a terrestrial planet, e.g. (**a**) in the case of Earth, a Mars-sized planetesimal, *Theia* is very likely to have led to the formation of the Moon; (**b**) the same process with different boundary conditions could result e.g. no re-accretion of mantle material, as it possibly occurred on Mercury early in its history

The giant impact on the Earth and comparable ones across other terrestrial planets, to a smaller extent contributed to the current compositions of mantles and to a larger extent to crusts of planetary bodies. The actual timing of the Moon-forming impact could have been relatively late, according to recent estimates up to 100 Myr after the initial formation of the Solar System.

11.2.3 Basin Formation

On the basis of Apollo and Luna samples a phase of intense bombardment at 3.9–4.0 Gyr has been proposed. Also known as the *Late Heavy Bombardment (LHB)* or the *Lunar Cataclysm*, most of the large basins on the Moon were supposedly formed during this time period. To explain this unusual spike in impact rate, outer Solar System dynamics have been suggested (Chap. 7 and excursion therein). According to this model, the giant planets and their gravitational interactions and resonances resulted in a perturbation of the asteroid belt and the Oort cloud (Chap. 1) to produce more projectiles to hit the Moon during this time period.

If it occurred, the Late Heavy Bombardment might have most likely affected all terrestrial planets with a comparable intensity, and, in the case of Earth, with

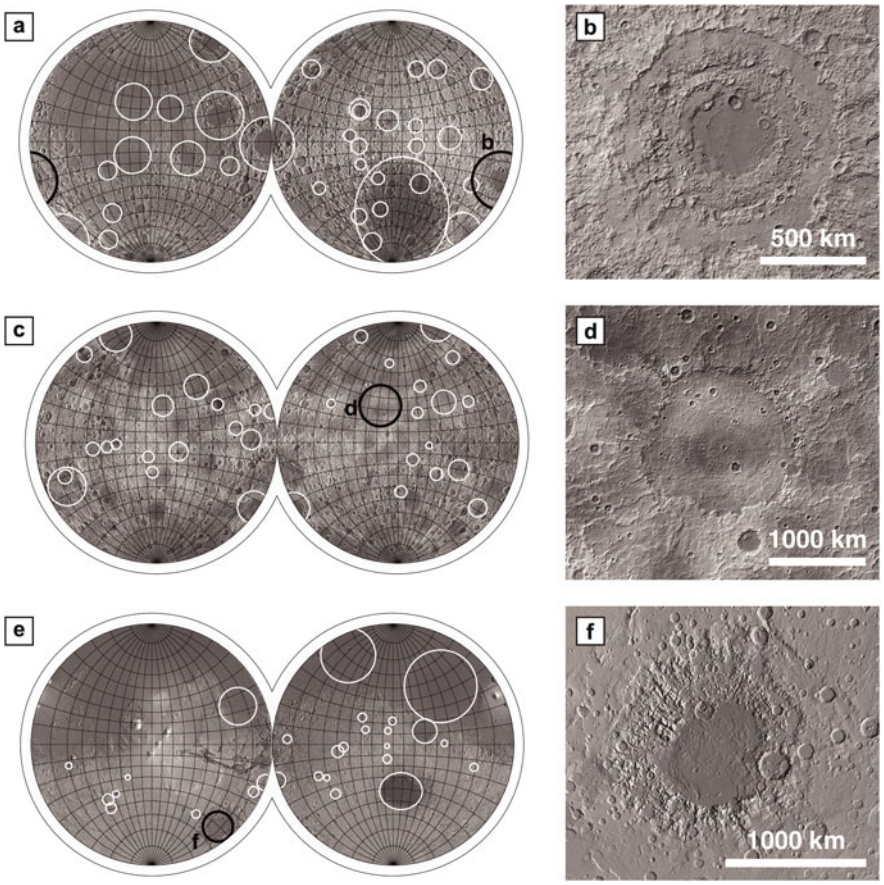

Fig. 11.9 A selection of large basins with various degrees of preservation and modification across the terrestrial planets. (**a**) For the Moon basins formed around 3.9 Gyr ago; (**b**) The lunar Orientale basin, an exemplary multi-ring impact basin of almost 1000 km diameter; (**c**) Mercury basins of ages close to that of the potential LHB; (**d**) The largest impact basin on Mercury, Caloris Planitia, has a diameter of about 700 km and less prominent rings when compared to lunar basins; (**e**) basins on Mars with ages comparable to that of the hypothesized LHB; (**f**) Argyre Planitia, 1800 km in diameter, appears more modified than similar counterparts on the Moon and Mercury, due to erosional and depositional processes. Sources: (**a**), (**c**), (**e**) Werner (2014). (**b**) NASA/LRO/LOLA. (**d**) NASA/MessengerMLA. (**f**) NASA/MGS/MOLA

possible partial or total sterilization (Chap. 14). This might have been the case also for the other terrestrial planets should life had been present back then.

Large impact basins, several hundreds to thousands of kilometres in diameter, are variably well preserved on the terrestrial planets (Fig. 11.9). If a *stratotype* of the LHB existed, the Moon would be the place for it: Large basins, well visible due to both lithological and structural differences to the surrounding primary anorthositic crust, are formed across a relatively short time span around 3.9 Gyr ago.

Mercury's basin are not as clearly outlined as those on the Moon and are different in morphology due to differences in gravity and internal structure. They have less well-developed rings and they are also more modified by global contraction occurring after their formation. On Mars, there are about 30 large basins of ages close to 4 Gyr (Fig. 11.9). They are less pristine than those on Mercury and the Moon, and this might be related to their strong post-impact modification either by volcanism or by erosional and depositional processes, such as ancient glacial activity potentially associated with Hellas and Argyre Planitiae. Consequently, partial or total obliteration of large basins on Mars might have occurred.

As with most other processes mentioned in this chapter, Venus' geological impact basin record is largely lost or not accessible by current data. So far, on Venus no large basins have been unambiguously detected, although one has been recently suggested based on contextual geological information. In fact, the largest confirmed impact crater, Meade, has a diameter of less than 300 km and is far younger than the age of a potential LHB on Venus. Also, the extensive later modification of Venus' surface might hinder the discovery of possibly heavily deformed craters in its oldest terrains.

Earth's geodynamics, dominated for most of its history by plate tectonics and its related crustal and lithospheric recycling, erased all morphological and structural signatures of any potential large impact structures. On Earth, plate tectonics itself possibly started with much smaller plates than those of today, due much more vigorous mantle convection (Chaps. 8 and 10). A possible trigger for early plate tectonics could have been the occurrence of large impact basins around ∼4 Gyr ago, close to the Hadean–Archean boundary (Fig. 11.4). The actual onset of modern plate tectonics is not well-constrained and estimates range from around 1 Gyr ago to about 4 Gyr, implying that it is potentially unrelated to the intense bombardment of the Hadean. The possible effects of the LHB on Earth today are only recorded in single crystals preserved in ancient terrains, after cycles of erosion and sedimentation have been modifying the original surface. The concepts of magma oceans, giant impacts and the LHB all condensed from detailed studies of lunar samples and remote sensing data as well as meteorites (Chap. 6). Without human or robotic sample return missions and access to the Moon, all those hypotheses could not have been developed in the same way.

Impact rates decreased through the history of the Solar System, although details are still debated (Chap. 7). Nevertheless, one of the cumulative effects of continued impact bombardment—both at large and small spatial/temporal scales—is the fragmentation of target rocks, which produces the other *regolith* layer. The regolith comprises the uppermost portion of planetary crusts, progressively and cumulatively disrupted by impacts of variable sizes, for geological timescales. Although the dominating process to form the Moon's regolith is impact-related physical disruption of existing material, other processes are active and are contributing to the uppermost portion of the regolith, also known as lunar soil. There are several definitions for the term *soil*, ranging in nature from granulometric to engineering-related properties: This is in contrast to the usage on Earth where it is mostly related to the development of an organic layer. The size range of individual

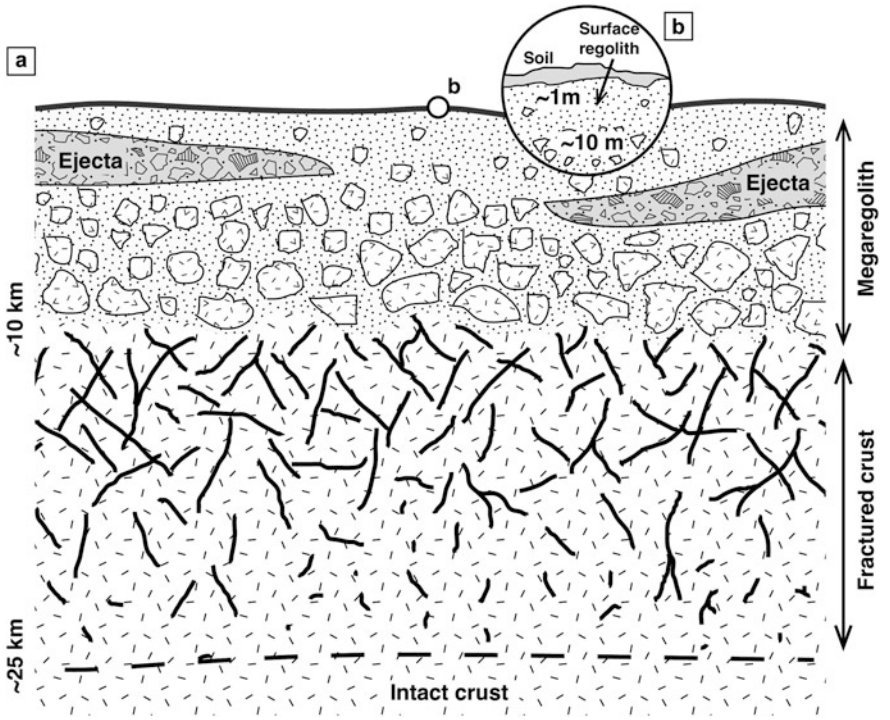

Fig. 11.10 Regolith and megaregolith development on planetary surfaces: (**a**) large-scale structure of the regolith/mega-regolith of the Moon. Regoliths/megaregoliths on other terrestrial planets, i.e., Mercury and to a lesser extent Mars, are also dominated for most of their geological history by impact cratering and, thus, should show similar characteristics; (**b**) enlargement of the uppermost portion of the crust, and the surface regolith. Source: after Hiesinger and Head (2006), Hörz et al. (1991)

components within the regolith and the megaregolith varies between microscopic (e.g., few micrometers) to extremely large, kilometre-sized or more (Fig. 11.10). In this respect, the fine-grained surface of planets is referred to as *regolith* whereas the heavily impact-disrupted upper crust down to tens of kilometers, is referred to as *megaregolith*. Regoliths progressively developed and they are the result of a cumulative set of processes, mainly driven by impacts. With the development and growth of the regolith, the underlying bedrock is increasingly protected from the effects of subsequent impacts craters smaller or comparable in size to that of the regolith thickness. Processes contributing to the progressive lunar soil formation (Fig. 11.10) include spallation, local fusion that cements granules into agglutinates as well as physical weathering due to temperature changes (Chap. 9), contributing to particle comminution specifically in the upper finer-grained portion of the regolith (soil).

On Venus the thick and dense atmosphere hindered small regolithforming impacts to reach the surface, while the record of megaregolith formation might have

been erased by the geologically recent vigorous resurfacing by mantle overturn. On Mars, the presence of an impact-generated, deep regolith (megaregolith) acted most likely as a subsurface reservoir for the martian hydrosphere. Without this regolith reservoir, the martian hydrosphere would have shrunk much earlier due to atmospheric loss by solar wind activity, in particularly as no significant magnetic field protected the planet for most of its history (Chap. 10). On the Earth, the preservation of a regolith or megaregolith layer has been prevented by its very dynamic crustal recycling as well as erosional and depositional processes acting throughout its geological history. Thus the terrestrial regolith, i.e., the soil, is produced by non-impact processes.

11.2.4 Secondary Crust Formation

After a period ranging from a few tens of Myr to a few Gyr and following the formation of an early crust and rigid lithosphere (Chap. 10), partial melting of the mantle and resulting basaltic eruptions produced a second type of crust on essentially all terrestrial planets (Chap. 8). This secondary crust is still very well preserved on the Moon, Mercury, and Mars at forms volcanic plains. For example, secondary crust constitutes the large basin-filling mare basalts on the Moon and may also be represented in large portions of Venus' surface. On Earth, secondary crust mostly builds up the ocean floors, hot spots, and other basalt plains. The accommodation space to host such secondary crusts for most terrestrial planets is provided by large impacts, producing multi-ring basins, in addition to the melt produced within the impact process itself (Chap. 7).

The duration of volcanic activity on the terrestrial planets is roughly scaled with their relative size and mass, and thus, their heat capacity. Small objects cool faster than larger objects due of their larger surface-to-volume ratio. Consequently, volcanism ceased rather early, possibly before ~3.5 Gyr on Mercury, below 3 Gyr and up to about 1.2 Gyr locally on the Moon (Fig. 11.11), and it lasted longer on Mars (most Amazonian units on Tharsis, as in Fig. 11.5).

Earth's modern secondary crust is mostly present on the ocean floors (Fig. 11.12). Historic ocean floors do not exist anymore as such as they have been consumed and closed by plate subduction and collision. On Earth, only few witnesses of even older, pre-plate tectonics crusts still exist. The remains are embedded in very old deformed terrains by processes active on Earth but presumably not on other terrestrial planets, with the exception of Venus. Secondary crust on the Earth was able to form even before the onset of modern plate tectonics, although the one preserved and formed today is linked to it.

The surface of Venus today is covered to more than 70% by volcanic plains (Fig. 11.12) that formed during a relatively recent, not well constrained time period roughly estimated somewhere between 100 Myr and ~1 Gyr. According to the model of *convective mantle overturn*, it was suggested that this global resurfacing event took place over a geologically short period (Chap. 8). It is possible that the

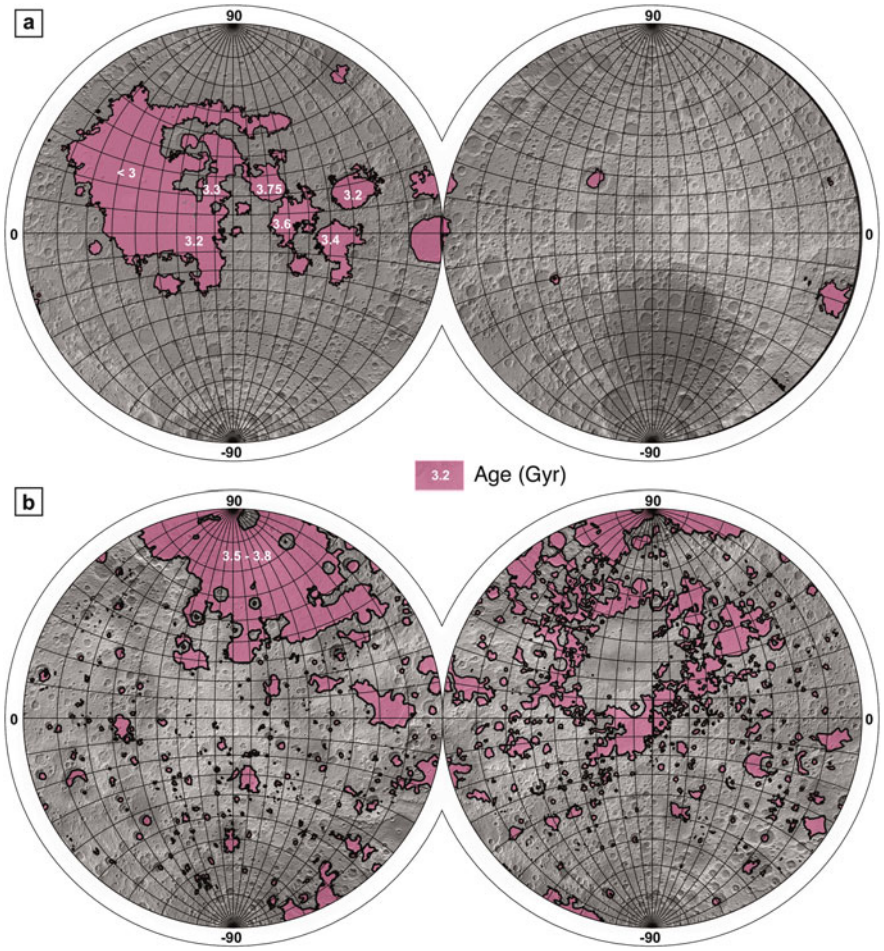

Fig. 11.11 Secondary crusts, composed of basaltic volcanic plains on terrestrial planetary bodies: (**a**) the Moon, see also sources of Fig. 8.12; (**b**) Mercury, smooth plains. Sources: (**a**) Fortezo and Hare (2013), see also sources in Fig. 8.12. (**b**) Procter et al. (2016), ages from Marchi et al. (2013)

event was not instantaneous or that several instances occurred. Nevertheless, some older terrains appear to have not been involved in such resurfacing and are still preserved, to an unknown extent. Today, these terrains are surrounded by more recent volcanic units with variable amounts of deformation (Chap. 8).

11.2.5 Continents and Planetary Counterparts

On Earth, the most evolved type of crust, i.e., continental crust, formed initially in relatively small volumes, which increased with time. It has been suggested that continental crust is only formed on planetary bodies that developed tectonic

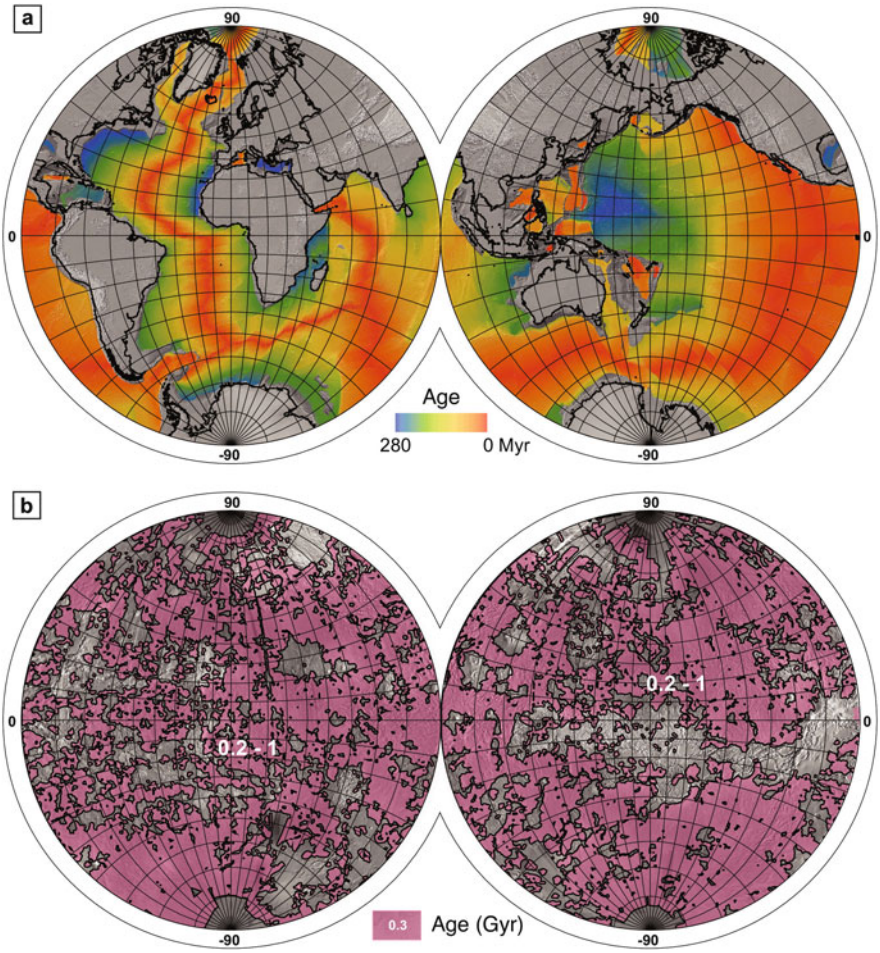

Fig. 11.12 Secondary crusts on the terrestrial planets: (**a**) Earth's recent crust, formed by partial melting of the mantle, is covering the oceanic floor. Older oceanic materials is recycled or embedded/obducted by plate tectonics and related mountain building; (**b**) Venus volcanic plains, covering about 70% of Venus' surface, have relatively young ages of up to several hundred million years, locally possibly much younger. Sources: (**a**) Müller et al. (2008), color-coded after Kovesi (2015). (**b**) Ivanov and Head (2011), courtesy M. Ivavov; age from Kreslavsky and Head (2015)

plates, i.e., Earth. However, on Mars, there is evidence of remnants of an ancient material compositionally comparable to Earths continental crust. Together with the formation of an evolved (tertiary) crust, this points to the possibility of an early onset of continental-like crust formation, which was possibly aborted after short time, and thus was not preserved in large geologically active bodies. Additional evidence of plate-like tectonics on Mars and large-scale horizontal movements such as those characteristic on Earth are ambiguous. Venus surface also does not provides unambiguous evidences for plate tectonics. Although individual geologic

and geodynamic features might be shared, locally, or regionally, they are not organized in a global framework (Chaps. 8 and 10).

Exhumation of old portions of the crust on the terrestrial planets is also achieved mainly through impact cratering, while on Earth plate convergence and collision allows for more efficient, although localised, exhumation of deep material, from several tens or even hundred (e.g. in large collisional chains) of kilometres of depth. Over geological timescales, eolian erosion can also cause significant exhumation, for example on Mars.

Venus displays few large continental-like terrains, called *tesserae*, (Fig. 11.13) that, among all terrestrial planets, are most similar to Earth's cratons. However, to date, no evidence (see Chap. 8) of comparable horizontal movement has been identified and accurate compositional information of the oldest terrains on Venus is lacking.

Given the different tectonic settings, mountain building among terrestrial planets and the Moon works very differently: on the Moon and Mercury local high relief is commonly produced by the formation of concentric rings associated with craters and large basins. Localized contraction is present on both Mars and, even more on Venus (Fig. 11.13). On the Earth, however, plate boundary interaction produces most topographic and structural relief as well as exhumation. Also, very ancient remains of terrestrial secondary crust are spread or squeezed into orogens with varying degree of preservation.

Orogenic and collisional processes, synchronously active in relatively small regions of the Earth's surface (i.e., *orogens*), are very effective in exhuming older or deeper portions of the crust, for example, by squeezing through or *obducting* crustal material. Over geological timescales, the areas affected by orogenic processes can grow and can form very large rock bodies (Fig. 11.13). Orogens create topographic highs subject to erosion that eventually expose deeper units and, at the same time, cause isostatic uplift due to the removed mass. Still, ancient orogens and their record are preserved and accessible. Once (for rocks from former oceanic domains) or since (for continental collision) these orogens are embedded into continental crust, they are less prone to be lost in the recycling of plate tectonics. On other terrestrial planets such recycling does not occur, although it could partly apply to Venus, associated possibly to the most intense localized compression.

On terrestrial planetary bodies, impact craters can serve as boreholes in the ground, exposing deep geological units. The original depth of such an exposed uplifted or ejected unit can be derived from measuring the size of the impact because its sampling (excavation) depth is well constrained by the crater size (Chap. 7).

On the terrestrial planets, the rock cycle might be affected in different ways on each planet, depending on its activity and geologic nature. In terms of basaltic production, the most common endogenic process, we observe significant spatial and temporal differences on the respective planets. Similarly, metamorphism, i.e., the solid-state changes of existing rocks can also be different. While products of impact metamorphism (Chaps. 6 and 7) are ubiquitous on the Moon, Mercury, and Mars because of the very large number of craters, they are restricted to the immediate vicinity of impact craters on Earth and Venus. Eolian sedimentary processes require

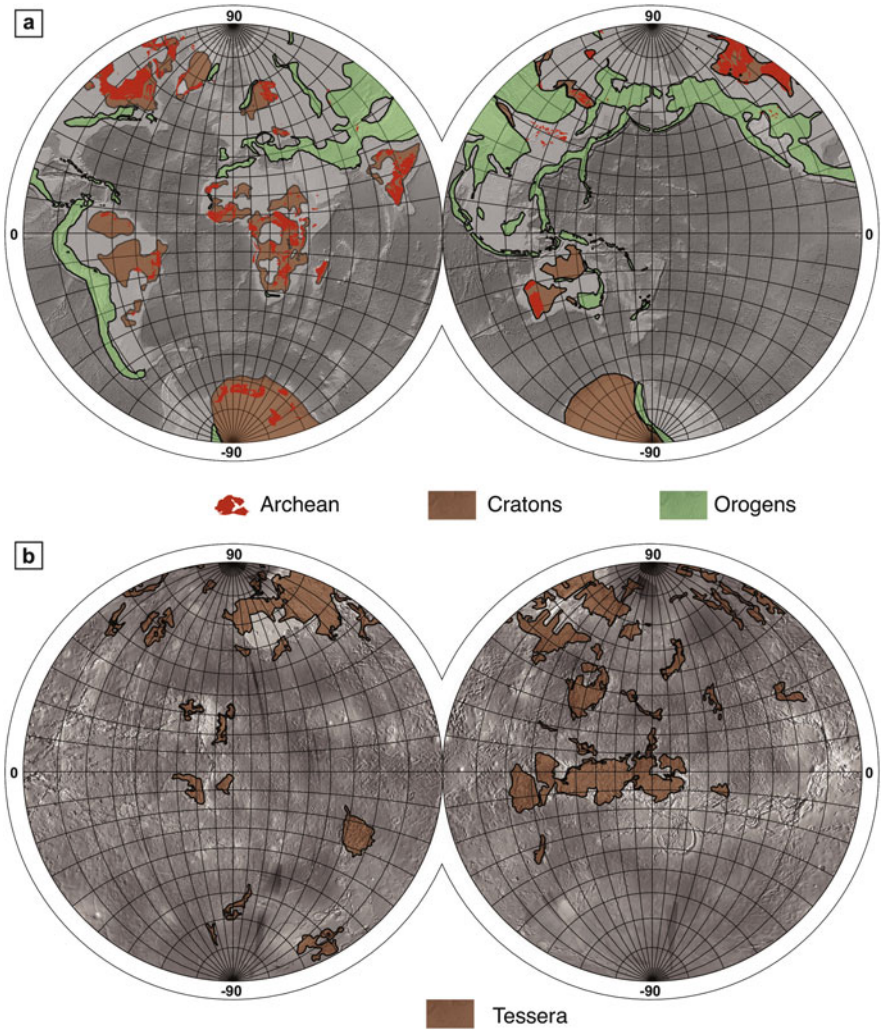

Fig. 11.13 Large-scale collisional features, orogens and cratons on Earth compared to Venus: (**a**) Global distribution of cratons and collisional orogens on Earth, as well as distribution of Archean rocks; (**b**) Venus' tessera terrain is highly deformed, it is older than the surrounding plain units, and occupies about 8% of the surface. Sources: (**a**) USGS. (**b**) Ivanov and Head (2011) courtesy M. Ivavov

a dynamic atmosphere and thus are mostly limited to Mars and Earth. Although Venus also has an atmosphere, the lack of sufficient sand-sized particles and observational effects of the radar data limit the detected number and mobility of sand dunes.

Physical and chemical weathering (Chap. 9) occurred both during different phases of the terrestrial planets' evolution. Today, chemical–physical weathering

is occurring on Venus surface and its highest mountain ranges might experience chemical weathering, resulting in the formation of radarbright peak areas. Mars' water-driven chemical weathering occurred mostly during the Noachian, while physical weathering continued throughout its history. On Earth, both kinds of weathering were constantly active for its entire history, mainly driven by climate. Climatic effects should be expected also on Mars, in addition to its long-term interior dynamics (Chap. 10), although their chronologic interplay is difficult to disentangle from their cumulative effect on the surface and near subsurface.

The evolution of sedimentary processes through time is very relevant also for future exploration, both for the possible link with life formation and their role in the creation of valuable resources (Chaps. 14 and 15).

11.2.6 Ancient Hydrologies and Surface Alteration

Among the terrestrial planets, Mars and Earth are the only planetary bodies that are located within the habitable zone. Thus, they have substantial amounts of water/ice that can interact and modify their surfaces. Most sedimentary processes (Chap. 9) are linked to an active hydrological cycle. Earth has such a cycle since very early in its history, most probably already during the Hadean and presumably even during periods of enhanced meteoritic bombardment, close to 4 Gyr ago. Throughout its history Earth most likely exhibited a surface that was covered at least to some extent by water.

Few billion years ago also Venus might have hosted a substantial amount of water on its surface and atmosphere, but the geological record of such possible ancient habitable Venus, suggested by recent climatic models, is unaccessible.

The Noachian period on Mars might bear resemblance to the earliest times of Earth's history, particularly before the onset of plate tectonics and life (Chap. 14). Even today, Mars resembles Earth in terms of processes that act on its surface, e.g., eolian, glacial, periglacial, volcanism, etc. (Chap. 2). However, the exact conditions during the Noachian are difficult to reconstruct without extensive ground truth and reliable proxies. There are several lines of evidences for water-related alteration and hydrated mineral formation, even pedogenesis, but the exact temperature and pressure ranges as well as the actual compositions are debated.

During the Noachian, Mars (Fig. 11.4) was characterized by a thicker atmosphere than today and, at least for few hundred million years, an internal magnetic field to protect the atmosphere (Chap. 10) from being eroded by large impacts, trapped in the subsurface or scavenged by the solar wind, though various processes. Atmospheric loss on Mars occurred for over 3.5 Gyr, and for comparable times also on Venus. However, Venus does not have an internal magnetic field to protect its atmosphere from escape but was able to replenish its atmosphere by outgassing associated with volcanism (Chaps. 8 and 10). Actual mechanisms leading to atmospheric escape by solar forcing are complex, ranging from thermal effects

to microscopic effects of the solar wind particles on atmospheric constituents, including photochemical escape, sputtering, ionisation, etc.

The complexity of the geological history of Mars and Venus is only second to that of Earth. In fact there are only two terrestrial planets with an ancient or present hydrological cycle and sedimentary processes, both erosional and depositional: Earth and Mars.

Mars' valley networks, comparable to Earth's drainage systems have commonly been interpreted to be linked to surface runoff and with variable contribution of groundwater (sapping, see Chap. 9) or surface ice-melting. They tend to be shallow, less than 100 m in depth and up to hundreds of km in length, with several networks reaching a few thousand kilometers in length. They are concentrated at low to mid-latitudes on Mars, typically in the highlands in the southern hemisphere, south of the dichotomy boundary (Fig. 11.14). Often, they are spatially and geologically associated with putative palaeolakes.

The D/H ratio measured on Venus has been interpreted to indicate substantial water loss in its past. If this loss of water did not occur too early, it might be plausible that Venus once hosted large bodies of water, developed a sedimentary cycle, and even might have been habitable (Chap. 14). However, this remains highly

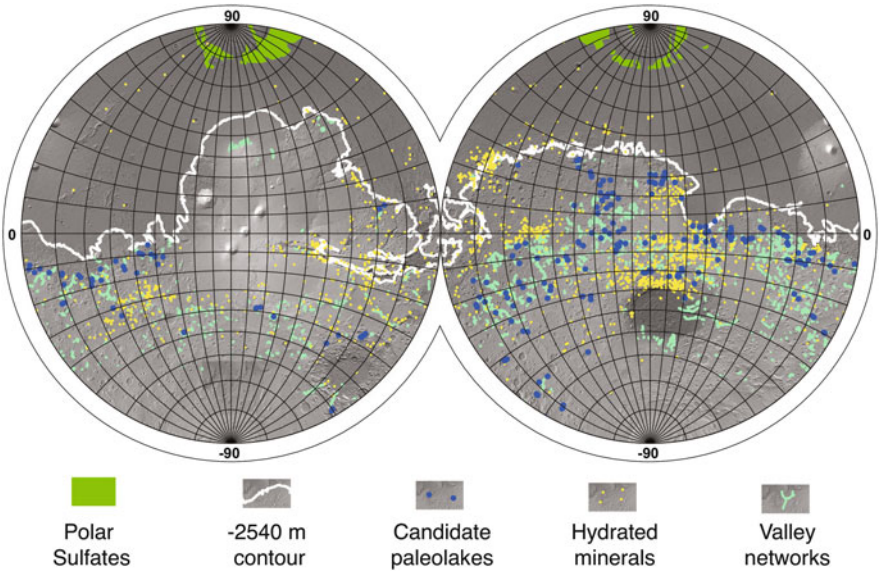

Fig. 11.14 Spatial distribution of selected geomorphologic features (paleolakes, valley networks) and mineralogical evidence (sulfates, hydrated minerals) in support of an ancient hydrosphere on Mars. *White line* is the −2540 m contour line, roughly demarcating the northern lowlands and the southern highlands. Source: Sulfate map courtesy J. Flahaut, after Massé et al. (2012); MGS MOLA contour after Di Achille and Hynek (2010); hydrated minerals compiled by Carter et al. (2013); Valley networks from Hynek et al. (2010); Open basin lakes from Fasset and Head (2008)

speculative because of the nature of crustal recycling on Venus, i.e., convective mantle overturn that erased all evidence for such scenarios.

On Earth, we have evidence for an ancient, stable, long-term sustained hydrosphere and related geological processes for at least the last 4 Gyr. For the same time period, Mars also shows evidence of relatively long-term liquid water stability at its surface, including erosional and depositional features such as valley networks, longitudinal valleys, outflow channels, deltas, paleolakes, or even putative temporary oceans as well as hydrated minerals (Chap. 9, see Fig. 11.14).

Sedimentary basins on Earth are mostly linked to plate tectonics (there are currently very few impact crater-hosted lakes, such as Bosumtwi in Ghana). However, in the Hadean and Archean, particularly before the onset of plate tectonics, it is likely that more impact basins existed on Earth. On Mars, numerous impact craters hosted palaeolakes as indicated by the deposits these paleolakes left behind. The exact estimate of the duration of such paleolakes is still a matter of debate and it ranges from transient to several Myr.

Geomorphologic evidence for water on Mars is abundant, has been recognized and investigated since the 1970s, and incudes large-scale fluvial and lacustrine features as well as glacial features, for example, in Argyre and Hellas Planitiae.

Additional evidence for water on Mars is provided by modern spacecraft that spectrally investigated the composition/mineralogy of the martian surface. As a result, numerous minerals were identified that were formed by water-related alteration of existing rocks (e.g., clay minerals) as well as primary deposition (e.g., carbonates). Although carbonate minerals have been detected in martian meteorites at a microscopic scale since decades (Chaps. 6 and 14), their orbital detection came later and is spatially limited to a few occurrences. Thus, despite the substantial amounts of water on Mars, widespread carbonate buildups like those on Earth never formed on Mars or were later destroyed. For example, perchlorates, as detected at the Phoenix landing site, might have hindered the formation of carbonates.

Many Noachian volcanic units on Mars (Fig. 11.5) display water-related alteration of their original basaltic composition, resulting in various types of phyllosilicates, including clays. Specific vertical distributions of those mineral associations have been detected regionally and have been interpreted as results of varying liquid compositions or the potential development of soil (*pedogenesis*).

On Mars, such alteration are limited to rather ancient times while they are common and ubiquitous on today's Earth. The absence of evidence for hydration in deposits later than Noachian can be explained, for example, by only a short-term intermittent presence of water that was unable to cause enough alteration in rocks. The exact environmental conditions for ancient Mars are still poorly constrained. For example, instead of the canonical *wet and warm* climate, early Mars could have been *cold and icy*, even if it was locally wet, and could have formed large-scale temperate glaciers (Chap. 9) with tremendous effects on surface alteration.

11.3 Intermediate, Diverging Histories

Interplanetary stratigraphies, although they follow similar patterns on the Moon and Mercury (Fig. 11.4), are not easy to correlate as geologic processes of similar nature acted asynchronously across the terrestrial planets. The timing, duration, and extent of volcanism, for example, is connected to the internal structure of a planet, its size, water content, and thermal characteristics, to name but a few. Similarly, the occurrence and nature of tectonism depend on the internal structure, the related stress field, the heat loss mechanism, and to some extent tidal deformation. Other processes are specific to planets that can hold an atmosphere, including fluvial, lacustrian, glacial, and eolian processes. Considering the common distant sources of the main impact-forming projectiles (i.e., the asteroid belt and the Oort cloud), it seems plausible that impact cratering might have affected the planets to a first order homogeneously. However, at closer inspection, the effects of the dense atmosphere on Venus, the distance of the planets from the Sun, different sizes and gravity, and many other factors affected the geological histories of the terrestrial planets.

Very early phases of complete surface layer melting, i.e., a magma ocean, and very intense bombardment by asteroids and comets are common to all terrestrial planets. The geological evolution of each planet then started to diverge, slowly or abruptly, depending on the studied processes. In the middle ages of planetary evolution, flood volcanism on the Moon and Mars, for example, continued to be active at different boundary conditions and eruption/depositional rates. During this phase, Mars and Earth divert progressively in terms of dominant active processes, although some of them act for comparably long times and even today (see also Chap. 2).

Although most geological evidence of liquid surface water is restricted to the Noachian or early Hesperian, some palaeolakes and valleys on Mars still occurred during the Hesperian. On Mars, the role of groundwater and subsurface hydrology became increasingly important during this phase as surface water became increasingly unstable.

11.3.1 From the Surface to the Subsurface

On Mars, the progressive loss of atmosphere and water over the last 4 Gyr, along with the associated reduction in atmospheric pressure and decrease in temperature, made liquid water unstable on the surface but stable in solid and vapor form. Beside atmospheric loss, water was also lost to the subsurface: Mars' crust, with its high porosity due to the impact-generated megaregolith (locally or regionally capped or sealed by other units, such as volcanic lava flows or sedimentary layers) acts as water reservoir (aquifer), of presumably local, regional, or even global scale. As temperatures dropped with time, the upper part of the water-saturated crust eventually became part of the cryosphere. Depending on, for example, the local

heat flow, heat conductivity, and surface temperatures, some models predict that the cryosphere sealed off possible liquid water beneath. Thus, it is plausible that there is still liquid water beneath the several kilometers thick martian cryosphere.

During the Noachian recharge of subsurface acquifers on Mars was achieved most likely through precipitation. As surface runoff became increasingly ineffective from most of the Hesperian onwards, aquifer recharge was also affected. Nevertheless the amount of subsurface water available when the Martian hydrological cycle stopped working, either in liquid or solid form, allowed for release of subsurface water at the surface through complex processes.

In addition to early lacustrine systems on Mars that have formed in crater basins, and large-scale surface runoff responsible for generating the global valley network distribution (Fig. 11.14) between the Noachian and the Hesperian, larger-scale catastrophic water outflow also occurred on Mars. These outflow channels possibly resulted in short-term ponding. The duration and extent of such ponding is still under debate and range from local deposits to hemispherical oceans. On the basis of absent large-scale water-related mineral alteration, both in distal and proximal portion of the large outflow features on Mars (Fig. 11.15), it has been argued that it is unlikely that the outflow channels produced long-lasting stable bodies of water on the surface.

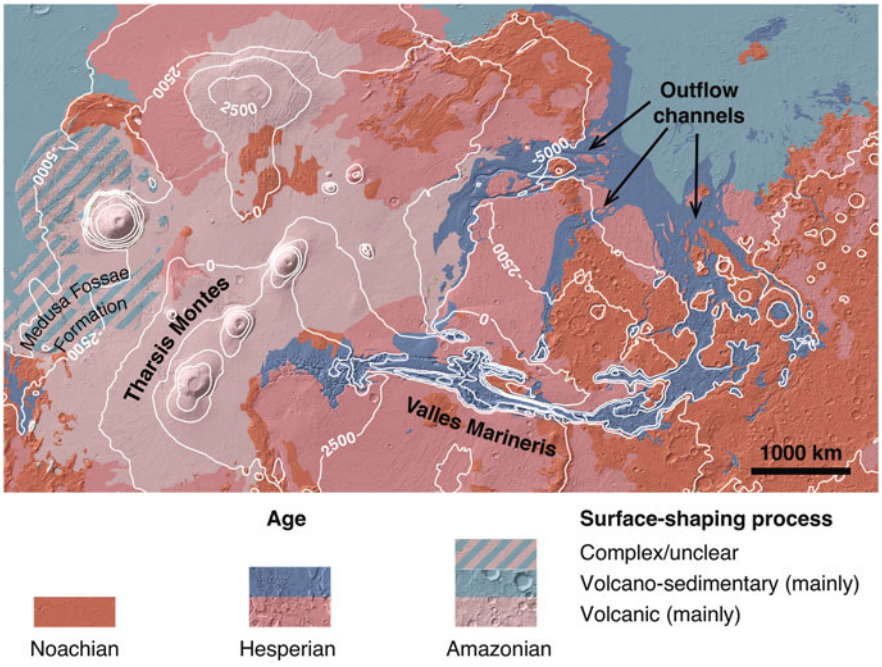

Fig. 11.15 Interaction of volcanic, tectonic and (catastrophic) sedimentary processes on Mars. Source: data adapted from Tanaka et al. (2014); design inspired by Carr (1996)

Interpreted in the literature as partially ice-carved or even volcanically-carved, outflow channels (Fig. 11.15) are largely considered to be the result of a catastrophic release of water from the subsurface, in one or more events. In this model, igneous processes or impact heating lead to melting of the cryosphere and the release of large amounts of subsurface water within short periods of time. Famous terrestrial analogues are known from Iceland (*jkulhlaups*) and Missoula, Washington. For example, the large, Tharsis province that was volcanically active for long times (Chap. 8) is spatially correlated with most outflow channels. Other smaller outflow channels are associated with regional (e.g., the circum-Hellas volcanic province) or local (e.g., Mangala Fossae) volcanic centers or dikes (Chap. 8).

Closely related to outflow channels are the *chaotic terrains*, located in many Hesperian and Amazonian craters along the dichotomy boundary (Fig. 11.15). Chaotic terrains vary in scale, from a few kilometers to hundreds of kilometers. In some cases it is linked to previous impact basins (e.g., Aram Chaos), in others it occurs as irregular depression, possibly due to the coalescence of several craters. Thus, in addition to heterogeneities within the megaregolith of the first several tens of kilometers of Mars' crust, the deformation within chaotic terrains is likely due to pre-existing impact-related structures (Chap. 7).

Several major outflow channel source areas are close to the Valles Marineris canyon system (Chap. 8). Some outflow channels were active only once, while others were active multiple times, as long as sufficient impact and/or volcanic heat and volatiles in the crust were present. Eventually Mars ran short of one or the other locally or regionally, thus, outflow channel activity did not last until the very recent geological past. Subsurface ice/water was not necessarily replenished, although volcanism might have continued later than the formation of outflow channels. In places where multiple episodes of outflow from roughly the same source occurred, the magnitude of floods in most cases became increasingly smaller, suggesting reservoir depletion with time.

Chaotic terrain and related liquid water-carved outflow channels are characteristic of Mars and have no counterpart in the inner Solar System, apart from some analogues on Earth that are associated with submarine slope failures. Some icy satellites display similar features, and were formed by disruption, however, in a very different geological context (Chap. 12). Chaotic terrain is widespread in a some regions on Mars possibly caused by the combined effect of low to midlatitude subsurface water release from the cryosphere in combination with volcanic, Tharsis-driven triggering of melting (Chap. 8). Although the occurrence of chaotic terrain appears to be spatially associated with outflow channels (Fig. 11.15), it exhibits various stage of development, from initial cracking to highly eroded mesa and knob morphologies. For a discussion of the role of collapse in Mars' chasmata and chaotic terrain see Chap. 8.

Since the shutdown of Mars' global magnetic field (Chap. 10) about 4 Gyr ago, the atmosphere of Mars has been continuously eroded by the solar wind. Such a change in global environmental conditions across the Noachian and Hesperian resulted in different geological/geochemical conditions and also in changes in dominant mineralogy. For example, widespread sulfates have only been formed

since the Hesperian—in some places in close vicinity of phyllosilicates—when available water became increasingly rare. Although these minerals can be reasonably well detected from spectrometers in orbit around Mars, landing missions will ultimatively provide ground truth for the occurrence of these minerals (Chap. 5).

The occurrence of sulfates has been interpreted as evidence for a very specific environment on Mars' surface and subsurface, i.e., more acidic conditions for most of its post-Noachian geological history. Some of the sulfates at low latitudes, for example those in Valles Marineris, are of Noachian to Hesperian age, and were either formed by primary or secondary processes. Others such as those in Terra Meridiani could be linked to evaporitic processes (Chap. 9). Finally, some are much younger and related to recent deposits, such as the circum-polar sulfate deposits linked to recent to active eolian landforms (Fig. 11.14). These sulfates possibly represent lag materials previously altered by and within ice rather than liquid water.

One large difference across terrestrial planets with an atmosphere (Venus, Earth, Mars) is the production of carbonates as local, regional, or global sinks of CO_2, either now or in the past. On Earth, life effectively supports the precipitation of carbonates to form extensive carbonate deposits. On Mars, the formation of carbonates is more local, at smaller scale, and most likely inorganic. On Venus, carbonates are unknown (Chap. 14).

11.3.2 Cryosphere and Water Loss

The existence of a past hydrological cycle on Mars is supported by several lines of geological evidence. On the basis of numerous studies it became clear that water was temporally stable at the surface, decreased in availability, some was lost to space, and some built up a partially global cryosphere.

The original amount of water on Mars surface has undergone large and difficult to quantify losses, with some models suggesting 95–99% loss of surface water. Today, some of this water is either stored in the subsurface cryosphere or was lost from the upper atmosphere. The many widespread outflow and collapse features, largely Hesperian and Amazonian in age, might have a common origin related to the release and recharge of subsurface water/ice, with an overall long-term volatile depletion. Mars' cryosphere has a much wider geographical and topographic distribution than on Earth, where the cryosphere is concentrated in higher latitudes and at high altitudes. Using radar sounding and neutron spectrometers, different spatial scales of the cryosphere could be explored, ranging from the first meter of regolith, as observed with neutron spectrometers to kilometer vertical scale as observed with orbital sounding radars (Chap. 3).

Some models predicted that the martian highlands were ice-covered for long periods of time during the Noachian, somewhat similar to the so-called snowball Earth state (Fig. 11.6). If correct, interaction between vigorous volcanism and such an ice cover could have produced liquid water early in Mars' history without puncturing the cryosphere as was the case to form the outflow channels mostly during the Hesperian.

Unlike geologically less diverse bodies such as the Moon and Mercury, Mars and Earth experienced, for most or all of their history, interactions between different geologic processes at multiple scales (Fig. 11.15). The interaction between distinct endogenic and exogenic geological processes through time produced the current surfaces on Mars and Earth (Fig. 11.17). The cumulative growth of continental crust on Earth in the last 4 Gyr and the growth of its younger oceanic crust are rather late events when compared to the surface ages on Mars. On dynamic Earth, geologic processes are still active and ongoing today whereas Mars' geologic activity is mostly ancient.

The progressive growth of continental crust and the plate tectonics-driven mobility of continents through time have a tremendous impact on climate, contributing to large-scale ice caps and global ice ages (Fig. 11.6). Those large-scale events, about 1.2 Gyr ago, are still preserved in the geologic record of Earth. Much shorter timescales are involved in recent ice ages, on Earth as well as on Mars.

Planet-wide or global climate changes occurred on three of the terrestrial planets, Venus, Earth, and Mars. These climate changes occurred at multiple temporal scales and are either linked to the large-scale configuration of lithospheric plates or driven by a variety of dominating processes, including orbital dynamics (Mars), endogenic activity or climate-endogenic coupling (Venus). On Earth, also effects of human activity are discussed to contribute to climate change (Fig. 11.16).

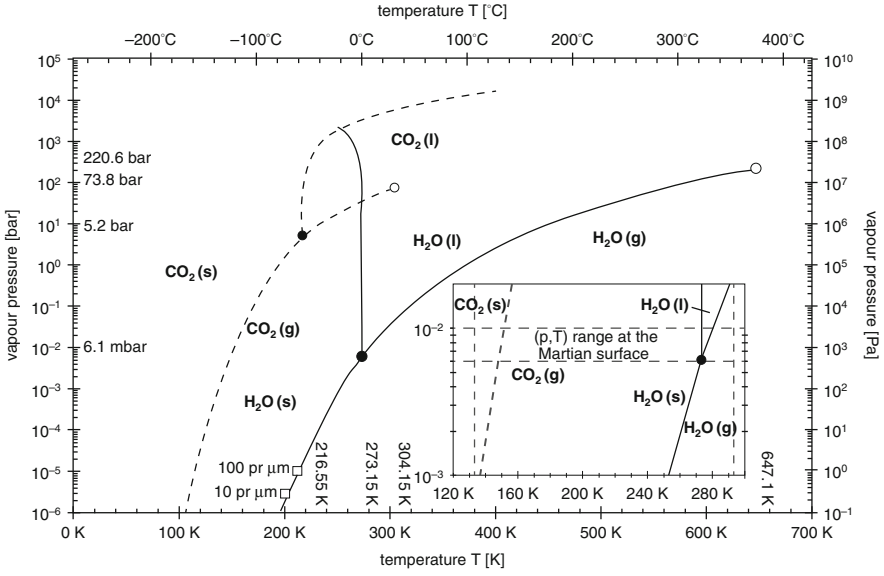

Fig. 11.16 Phase diagram for water and carbon dioxide as well as frost-point temperatures for a well-mixed atmosphere and 10 pr µm and 100 pr µm atmospheric water vapor. *Filled and empty circles* are respectively critical and triple points. Source: figure from van Gasselt (2007) and reference therein

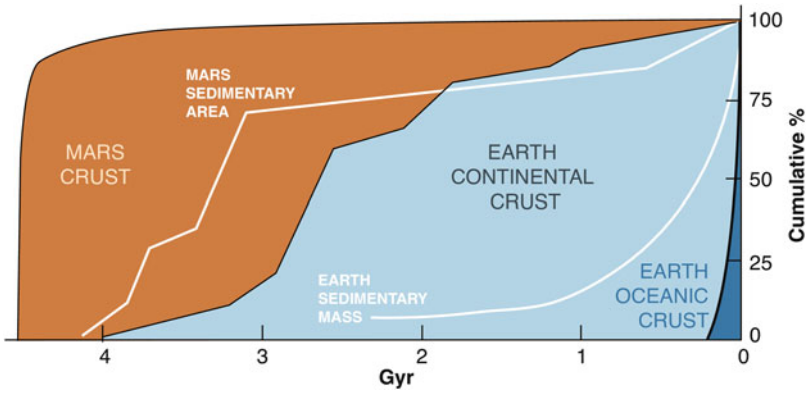

Fig. 11.17 Relative comparison of Earth's and Mars' crust production through times and sedimentary rocks. Volumetric information available on Earth is lacking on Mars, thus surface area as measured on global geological maps is used. Source: modified from McLennan (2012), after Taylor and McLennan (2009)

Globally acting sedimentary processes as recorded by extensive sedimentary deposits are preserved and more or less accessible only on Earth and Mars. Although Venus might host ancient sedimentary deposits as well, they still need to be identified. On Earth and Mars, the amounts of sedimentary rocks produced through time follow the respective degree of preserved geological activity (Fig. 11.17). Mars, for example, had an early intense period of global geological activity followed by less intense regional and local geological processes. While volcanism lasted quite long, until the recent geological past (~10–100 Myr), most sedimentary rock formation dates back to the Noachian/Hesperian, with far less steep production rates in the Amazonian. Earth's sedimentary record is best preserved in the Phanerozoic, while its present oceanic basins are of even younger age (Fig. 11.12).

11.4 Recent Phases

Each planet has seen its own geological evolution and the timing of geological processes is not necessarily contemporaneous among these objects. On Earth, the last billion years make up a substantial part of its recorded history, during which endogenic activity (Chap. 8), life (Chap. 14), modern sedimentary processes (Chaps. 2 and 9), and most modern plate tectonics processes occurred. In contrast, for the Moon, the last billion years is a period of extremely low activity, characterized by continued impact bombardment and small-scale volcanic activity. Similarly, Mercury also shows evidence for low activity during this time, although some late-stage local volcanic processes have been discovered on its surface.

Mars on the other hand, has experienced quite some geological activity, although not comparable with the activity taking place in the Noachian and Hesperian. For example, most of the volcanism linked to the Tharsis province is Amazonian in age (Fig. 11.4), and the observed polar caps are also considered to be recent features.

Large-scale plate tectonics dynamics have been linked to long-term climatic conditions on Earth. The assemblage of supercontinents occurred several times (Fig. 11.6) since the Proterozoic (e.g. Rhodinia, Pangaea) and was associated with large-scale continental glaciations (Fig. 11.6). Possibly, similar phases of glaciation also occurred much earlier, i.e., associated with older supercontinents.

On a different time scale, of few million years or less, Quaternary ice ages on Earth are globally recorded in landforms and the geological record, e.g. in proxies from sediment cores.

The absolute magnitude of astronomical variations leading to terrestrial ice ages is relatively small. The tidal locking and synchronous rotation (Chap. 1) of the Earth-Moon system stabilizes the inclination of Earth's spin axis obliquity, that varies only within little more than two degrees. On Mars, due to the lack of sizeable moons, such a stabilizing effect is not present and the range of variation of its obliquity over timescales of few million years can be of several tens of degrees. Such variations had large impacts on Mars' cryosphere and overall volatile distribution in the subsurface, surface, and atmosphere. The location of polar caps during late Amazonian ice ages on Mars could have been in different geographic positions. However, because the locations of potential palaeo-polar caps are not well established, there exist different interpretations of specific widespread geological units, such as the Medusa Fossae Formation (Fig. 11.15), ranging from volcaniclastic to glacial.

During higher obliquity periods, larger amounts of volatiles were injected into Mars' atmosphere, possibly resulting in snow precipitation, at least locally. Such short term (few million years) snow accumulation could produce glaciers, whose possible geomorphologic traces have been identified on Mars in several locations (Chap. 9). Evidence of widespread recent periglacial features on Mars also exist. Permafrostrelated features are very common in the mid and high latitudes on Mars. In particular, features such as gullies, which on Earth are often linked often to near subsurface ice melting or precipitation of snow, are widespread in higher mid latitudes. Although their exact mechanism of formation is debated, they were active in the very recent geological past, possibly during or after higher obliquity phases with increased subsurface ice melting. Alternatively, dry processes or liquid water sources have also been suggested for explaining their development.

On Mars, chasmata and canyon development continued locally into the Amazonian. It is also plausible that, volatiles were exchanged between the polar caps, the ice in the regolith, and the atmosphere. The dynamics of such exchange processes were also linked to climatic variations. Recent processes on Mars include the formation of perchlorates, very reactive compounds detected for the first time on the northern plains at the Phoenix landing site, (Chap. 15). In the presence of ice, these perchlorates could result in brines formed near the surface (Chap. 9).

It has been proposed that brines might be responsible for the formation of recurring slope lineations (RSL), i.e., flow-like surface changes on sloped surfaces. These RSL features have been detected in different locations and are spatially associated with the presence of hydrated salts. Although possible sources of water responsible for such flow features on present Mars are not clear (subsurface, atmosphere), their current state of activity has been well-documented.

Aeolian processes occur on three out of four terrestrial planets (Earth, Mars, and Venus), and are also active on some icy satellites (Chap. 12), and might even affect the surface on comets (Chap. 13). Active depositional features are documented on Mars and are likely on Venus. Unlike on Earth where aeolian sand tends to be largely made of quartz and igneous minerals are rather rapidly weathered, Mars' dunes are mostly composed of basaltic sand. On Earth, deserts and sand seas occur predominantly at relatively low latitudes; on Mars there is a clustering of dune fields close to the north polar cap (Fig. 11.18). On Venus, aeolian depositional features were not yet imaged in enough detail to resolve their characteristics.

Surface-atmosphere interaction is important on all major terrestrial planets, including presentday weathering. In particular, transient and local, but very widespread phenomena such as dust devils are responsible for large amounts of dust injected into the lower atmosphere of Mars, with possible effects on the climate.

In the last billion years, little geologic activity occurred on Mercury and the Moon. Such activity is mostly limited to impacts (e.g., the formation of the lunar Copernicus crater in the last few hundred Myr), local mass wasting, and the effects of space weathering on the surface regolith, including implantation of ions from the solar wind (Chap. 15). Mass wasting driven by gravity, impacts and possible internal activity is an ongoing process occurring at low pace, on several terrestrial planetary bodies, such as Mars and the Moon. Recent mass wasting processes were imaged on the Moon, while active avalanches on polar deposits were imaged on Mars by orbiters several times. On the other hand, in the very same last billion years or so Venus experienced most of its recorded geological activity. The basis for reconstructing Venus' global history and evolution is largely geological mapping (Chap. 4), but the integrated study of its climate and atmospheric evolution, including modelling, is important too (Chap. 11).

On Venus, the earliest unit preserved, the tessera terrain, covers approximately 8% of Venu's surface (Fig. 11.13), and recorded early large-scale contractional and extensional phases, likely during geologically short, very intense deformation phases. Substantial deformation also affected the earliest post-tessera volcanic plains. This deformation was concentrated locally and resulted in ridge belts (Chap. 8). On Venus, the detailed geological history can only be described in a relative sense, not chronostratigraphically, because the homogeneous and random distribution of impact craters does not allow dating of specific surface units. Regional to global events of extension and compression alternate in the youngest units and localized extension associated with volcanic rises (e.g., similar to what occurs on Earth) is present. Overall, it appears that volcanic and tectonic processes recorded on Venus' surface were active globally. Similar to other terrestrial planets other than Earth, their magnitude appears to have decreased with time.

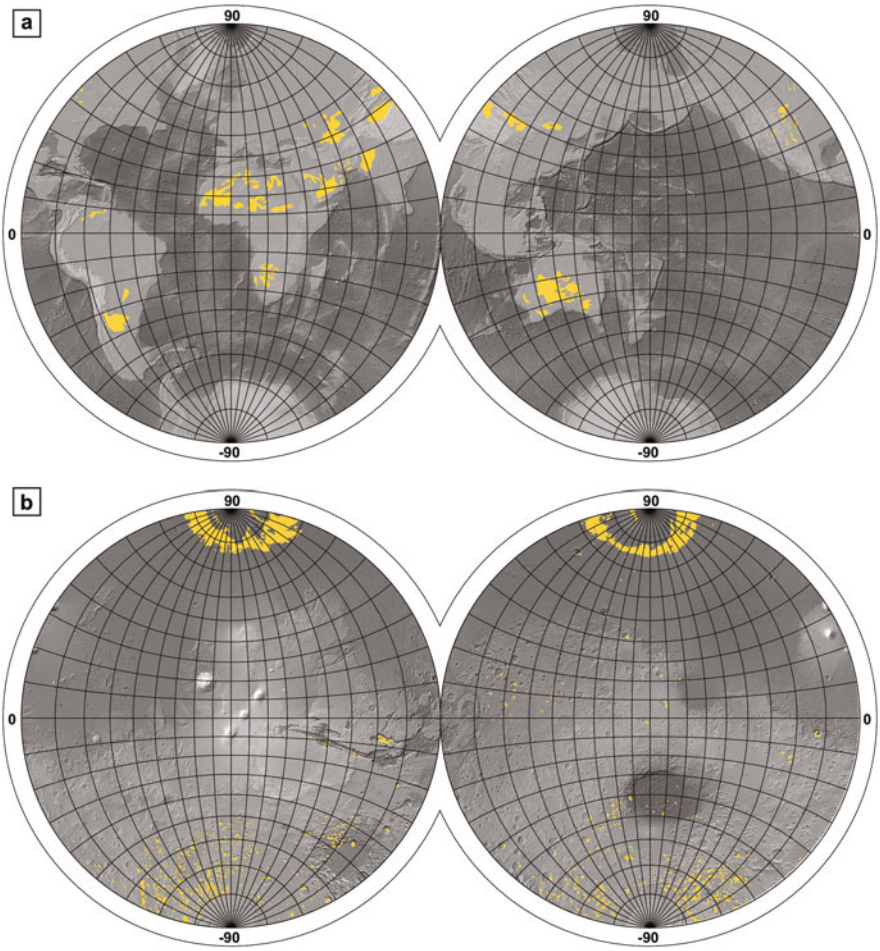

Fig. 11.18 Aeolian processes and products are present on Earth, Mars and Venus. The largest deposits are present on Earth and Mars; (**a**) global map of sand seas on Earth; (**b**) global occurrence of sand dunes on Mars. Sources: (**a**) data from Sun and Muhs (2007). (**b**) Hayward et al. (2007, 2010, 2012)

In its recent geological past (less than 1 Gyr), Venus experienced a more or less global volcanic resurfacing that erased almost all existing terrains as well as the evidence for the early geologic history. The coupled endogenic and atmospheric evolution produced a very pronounced climate change. Although ancient, pre-resurfacing surface conditions are not accessible anymore, the isotopic record of the atmosphere and numerical climatic models hint at possibly very different conditions, compared to today's greenhouse.

The range of documented active processes on the terrestrial planets is wide (Table 11.1): space weathering and continued impact cratering are the most

Table 11.1 Comparison of recent and current processes on the terrestrial planets and the Moon

Body	Endogenic	Surface	Surface/atmosphere
Mercury	None	Space weathering, volatile escape/exchange in polar areas	Loss to exosphere
Venus	Recent, possibly active volcanism	Aeolian processes, surface physical and chemical alteration	
Earth	Plate Tectonics, volcanism	Active hydrology and sedimentary cycle	Ample surface-atmosphere interaction, alteration, pedogenisation
Moon	Low seismic activity, recent volcanism?	Space weathering, mass wasting	Volatile escape/exchange in polar areas
Mars	Recent Volcanism, possibly recent/present seismic activity (?)	Active aeolian processes	Volatile exchange between polar caps, regolith and atmosphere, possible local surface brines

Impact processes are not included because they are ubiquitous on all planets with the exception of Venus, where most small impactors never reach the surface because of the thick atmosphere (see Chaps. 8 and 10)

important processes on virtually airless bodies (the Moon, Mercury) while surface-interior and surface-atmospheric interaction is important on other planets (Venus, Earth, Mars).

11.4.1 Planetary Global Change and Perspectives

Planetary sciences and geology traditionally address long-term and large-scale evolutionary aspects of planets, although uncertainties are often high, particularly for targets or processes observed for the first time (Chap. 1). Data of planets such as Mars have by now excellent spatial and temporal coverage, although not always continuous.

One of the messages conveyed by the cumulated geological understanding of the terrestrial planets is that both internal and external dynamics globally affect the planets. Examples of the latter include increased phases of impact cratering (e.g., LHB), related to both enhanced atmospheric erosion and surface disruption or subsurface mobilization. Internal processes such as global partial melting of the mantle and resurfacing on Venus as well as plate tectonics on Earth have/had profound, long-lived global consequences for the entire planet. In Earth's case, the onset of life triggered changes both in the geosphere and atmosphere (Chap. 14).

Recent to current processes across terrestrial planets that bear certain resemblance, although at largely different spatial and temporal scales, include, for example: The greenhouse effect (Venus, Earth), and recent climatic cycles (Mars, Earth). The extreme climatic conditions and greenhouse effect on Venus are mainly linked to its geodynamics. On Earth, the change from a CO_2-rich atmosphere to a nitrogen/oxygen atmosphere is related to the evolution of life. Recent, non negligible effects on the geologic evolution of Earth are linked to human activity. Thus, the formalization of anthropogenic effects in the geological records has been discussed lately and might eventually result in a formal stratigraphic epoch name (Chap. 2). On Mars, recent climate cycles and their effects on surface landforms are well-recorded and are similar in scale and processes to those on present glacial or periglacial areas on Earth, particularly to selected, very cold and very dry locations such as the Dry Valleys in Antarctica (Chaps. 2 and 9).

As much as the traces of human activities would most likely be retained in Earth's geological record, human artifact on other planets such as spacecrafts, structures and tools left behind on the Moon (Chap. 5), are likely to be preserved for geological timescales.

Speculating on the geological future of terrestrial planets is less difficult for the Moon and Mercury compared to the other terrestrial objects. Because the Moon and Mercury are small bodies that already cooled to large extents, not much internal geological activity is expected to take place on them. Venus, is characterized by episodic convective mantle overturn that recycles the surface at global scales. Although little to no volcanic activity is currently observed, at time scales of hundreds of million years or billion years, it is expected that Venus will experience periods of extreme activity. On Earth, geologically long-term changes in the tectonic regime, such as a stop of plate tectonics and an evolution toward a stagnant lid regime (Chap. 10) could be a few billion years ahead. Thus, overall changes in the atmosphere and the near subsurface, for example, with the release of volatiles from ground or seabed reservoirs (e.g., chlathrates) are more likely in the geological near term. The effects of those changes might be substantial and even tragic for mankind. The related geological record is likely to be condensed just on relatively thin deposits, although globally correlated.

The understanding of Earth-like planets such as Venus and Mars can shed light on both past and future processes, whose record is incomplete or absent on Earth. The Moon is unique in that it is history book of more than 4.5 Gyr Solar System evolution which retains a geological record that is very complementary to ours. The growing understanding of the most remote terrestrial planet, Mercury, can inform us on how Earth evolved compositionally, e.g., with respect to a late delivery of carbon by a C-rich planetesimal, with similarity to Mercury, also more C-rich, compared to other terrestrial planets. The growing evidence from astronomical observations of terrestrial exoplanets enormously widens the range of possible geological boundary conditions; the most accessible examples though, are those in the Solar System.

Take-Home Messages

Terrestrial planets share several characteristics, including a variably preserved impact cratering history, silicate volcanism and tectonic deformation driven by very diverse geodynamics.

The bulk of the geological activity is concentrated at different age ranges for different planets: intense and violent early phases for all planets, followed by variable phases of volcanism, deformation and, for Mars and Earth, sedimentary activity.

The Moon, Mercury, and to a large extent also Mars allow access to the geological record of the early Solar System. Earth and Venus retain little information about the first hundred Myr of Solar System history but have a better record of the most recent several hundred Myr.

Beside impact cratering, volcanism and tectonism dominated the geological evolution of all terrestrial planets and the Moon, particularly during the first 1–2.5 Gyr. Mars exhibits large-scale volcanic activity that continued past the first 2.5 Gyr. Venus and Earth are the two planets that were most recently/currently volcanically active.

Sedimentary processes linked to a transient or stable hydrosphere are recorded on Earth and Mars. Venus could have hosted a hydrological cycle during its early phases of geologic evolution, but evidence for such a cycle has been erased from its geological record.

The level of current activity is variable on the terrestrial planets. Apart from the present low impact rate, common to all planets, Mercury and the Moon are the least active. Venus is possibly active both volcanically and tectonically. On Mars and Earth numerous geologic processes are active today.

Global, often catastrophic changes are documented, within different geological boundary conditions, on all terrestrial planets.

Further Readings

Arvidson, R.: Aqueous history of Mars as inferred from landed mission measurements of rocks, soils, and water ice. J. Geophys. Res. Planets. **121**(9), 1602–1626 (2016). doi:10.1002/2016JE005079.

Balme, M., Gallagher, C., Hauber, E.: Morphological evidence for geologically young thaw of ice on Mars: a review of recent studies using high-resolution imaging data. Prog. Phys. Geogr. **37**(3), 289–324 (2013). doi:10.1177/0309133313477123.

Basilevsky, A.T., Head, J.W.: The geologic history of Venus: a stratigraphic view. J. Geophys. Res. Planets **103**(E4), 8531–8544 (1998). doi:10.1029/98JE00487

Carr, M.H.: Channels and valleys on Mars: cold climate features formed as a result of a thickening cryosphere. Planet. Space Sci. **44**(11), 1411–1423 (1996). doi:10.1016/S0032–0633(96)00053-0

Carr, M.H., Head, J.W.: Geologic history of Mars. Earth Planet. Sci. Lett. **294**(3), 185–203 (2010). doi:10.1016/j.epsl.2009.06.042

Carr, M.H., Head, J.: Martian surface/near-surface water inventory: sources, sinks, and changes with time. Geophys. Res. Lett. **42**(3), 726–732 (2015). doi:10.1002/2014GL062464

Ehlmann, B.L., Edwards, C.S.: Mineralogy of the Martian surface. Annu. Rev. Earth Planet. Sci. **42**(1), 291–315 (2014). doi:10.1146/annurev-earth-060313-055024

Elkins-Tanton, L.T.: Magma oceans in the inner solar system. Annu. Rev. Earth Planet. Sci. **40**(1), 113–139 (2012). doi:10.1146/annurev-earth-042711-105503

Farley, K.A., et al.: In situ radiometric and exposure age dating of the martian surface. Science **343**(6169), 1247166 (2014). doi:10.1126/science.1247166

Geiss, J., Rossi, A.P.: On the chronology of lunar origin and evolution. Astron. Astrophys. Rev. **21**(1), 1–54 (2013). doi:10.1007/s00159-013-0068-1

Gregg, T.: (2015). Planetary tectonics and volcanism: the inner solar system. In: Schubert, G. (ed.) Physics of Terrestrial Planets and Moons. Treatise on Geophysics, vol. 10, Chap. 9, pp. 307–325, 2nd edn. Elsevier, Oxford (2015). doi:10.1016/B978-0-444-53802-4.00187-1

Grotzinger, J., Hayes, A., Lamb, M., McLennan, S.: (2013). Sedimentary processes on Earth, Mars, Titan, and Venus. Comp. Climatol. Terr. Planets **1**, 439–472 (2013)

Hansen, V.L.: Impact origin of Archean cratons. Lithosphere **7**(5), 563–578 (2015). doi:10.1130/L371.1

Hansen, V.L., Young, D.: Venus's evolution: a synthesis. Geol. Soc. Am. Spec. Pap. **419**, 255–273 (2007). doi:10.1130/2006.2419(13)

Hartmann, W.K.: The giant impact hypothesis: past, present (and future?). Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci. **372**(2024) (2014)

Helbert, J., Hauber, E., Reiss, D.: Water on the terrestrial planets. In: Schubert, G. (ed.) Physics of Terrestrial Planets and Moons. Treatise on Geophysics, vol. 10, Chap. 11, pp. 367–409, 2nd edn. Elsevier, Oxford (2015). doi:10.1016/B978–0-444-53802-4.00174-3

Hiesinger, H., Head, J.W.: New views of lunar geoscience: an introduction and overview. Rev. Mineral. Geochem. **60**(1), 1–81 (2006). doi:10.2138/rmg.2006.60.1

Lasue, J., Mangold, N., Hauber, E., Clifford, S., Feldman, W., Gasnault, O., Grima, C., Maurice, S., Mousis, O.: Quantitative assessments of the martian hydrosphere. Space Sci. Rev. **174**(1), 155–212 (2013). doi:10.1007/s11214-012-9946-5

McLennan, S.M., et al.: Geochemistry of sedimentary processes on Mars. Sediment. Geol. Mars **102**, 119–138 (2012)

Reiss, D., Lorenz, R., Balme, M., Neakrase, L., Rossi, A.P., Spiga, A., Zarnecki, J. (eds.) Dust Devils. Space Sciences Series of ISSI, vol. 59, 426 pp. Springer (2017). ISBN: 978-94-024-1133-1, ISSN: 1385-7525

Smrekar, S.E., Stofan, E.R., Mueller, N., Treiman, A., Elkins-Tanton, L., Helbert, J., Piccioni, G., Drossart, P.: Recent hotspot volcanism on Venus from VIRTIS emissivity data. Science **328**(5978), 605–608 (2010). doi:10.1126/science.1186785

Steffen, W., Leinfelder, R., Zalasiewicz, J., Waters, C.N., Williams, M., Summerhayes, C., Barnosky, A.D., Cearreta, A., Crutzen, P., Edgeworth, M., Ellis, E.C., Fairchild, I.J., Galuszka, A., Grinevald, J., Haywood, A., Ivar do Sul, J., Jeandel, C., McNeill, J., Odada, E., Oreskes, N., Revkin, A., Richter, D.D., Syvitski, J., Vidas, D., Wagreich, M., Wing, S.L., Wolfe, A.P., Schellnhuber, H.: Stratigraphic and Earth System approaches to defining the Anthropocene. Earth's Futur. (2016). doi:10.1002/2016EF000379

Taylor, S.R., McLennan, S.: Planetary Crusts: Their Composition, Origin and Evolution, vol. 10. Cambridge University Press, Cambridge (2009)

Way, M.J., Del Genio, A.D., Kiang, N.Y., Sohl, L.E., Grinspoon, D.H., Aleinov, I., Kelley, M., Clune, T.: Was Venus the first habitable world of our solar system? Geophys. Res. Lett. (2016). doi:10.1002/2016GL069790

Werner, S.C., Ody, A., Poulet, F.: The source crater of Martian Shergottite meteorites. Science **343**(6177), 1343–1346 (2014). doi:10.1126/science.1247282

Wilhelms, D.E., John, F., Trask, N.J.: The geologic history of the Moon. USGS Professional Paper, vol. 1348. U.S. Geological Survey. http://ser.sese.asu.edu/GHM (1987)

Wordsworth, R.D., Kerber, L., Pierrehumbert, R.T., Forget, F., Head, J.W.: Comparison of warm and wetİ and cold and icyİ scenarios for early Mars in a 3-D climate model. J. Geophys. Res. Planets **120**(6), 1201–1219 (2015). doi:10.1002/2015JE004787

Chapter 12
Icy and Rocky–Icy Satellites

Roland Wagner, Katrin Stephan, and Nico Schmedemann

12.1 The Icy and Rocky–Icy Satellites of Jupiter

The largest planet of our solar system, Jupiter, has a total of 67 satellites that are currently known, most of them bodies much smaller than 250 km in diameter (see Chap. 13). The four largest moons were confirmed to be observed for the first time by GALILEO GALILEI and SIMON MARIUS in 1609/1610 and termed *Galilean Satellites*. The two largest ones, Callisto and Ganymede, are about the size of the planet Mercury while Europa and Io are about the size of the Earth's moon. Io is a rocky terrestrial-type planetary body consisting mostly of metal and silicates (see excursion on Io). These moons are thought to orbit Jupiter with synchronous rotation like the Earth's moon, but non-synchronous rotation is in discussion. Primary characteristics of these four satellites are an increase in geologic activity and a decreasing surface age with decreasing distance towards Jupiter from Callisto to Io—a strong trend not observed in other satellite systems. After two flybys of Voyager 1 and 2 in 1979, Jupiter and its satellites were investigated in detail by the remote sensing instruments aboard the Galileo Orbiter between 1995 and 2003, including the high-resolution camera experiment *Solid State Imaging (SSI)* and the *Near-Infrared Mapping Spectrometer (NIMS)*.

For icy satellite geology in general, a new term was introduced known as cryovolcanism. Cryovolcanic processes include the extrusion of liquid, gaseous or solid-particle materials, or a mixture of these, as in the case of volcanism on terrestrial planets. Unlike molten silicates and volatiles on these planets, liquid, solid

R. Wagner (✉) • K. Stephan
German Aerospace Center (DLR), Berlin, Germany
e-mail: roland.wagner@dlr.de; katrin.stephan@dlr.de

N. Schmedemann
Freie Universität Berlin, Berlin, Germany
e-mail: nico.schmedemann@fu-berlin.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_12

and gaseous volatile species such as H_2O, CO, CO_2, or organic materials erupt on icy satellites (or icy dwarf planets, see Chap. 13) and are emplaced on the surface and/or ejected from the body.

12.1.1 The Callisto–Ganymede Dichotomy

Callisto (4818 km diameter, 1834 kg/m^3 mean density) and Ganymede (5264 km diameter, 1942 kg/m^3 mean density) are nearly the same size but their surfaces are remarkably different suggesting they have followed two completely different evolutionary paths during their histories, a feature commonly termed the Ganymede–Callisto dichotomy. Global as well as detailed views of the two satellites highlight this difference (Figs. 12.1 and 12.2).

Both satellites have comparably high albedos. Callisto has the lowest albedo ($\sim$0.11) of the Galilean satellites but is still twice as bright as the Earth's moon. Ganymede's average albedo is $\sim$0.43 but varies with terrain type. The abundance of various ices which were spectroscopically detected, such as predominantly H_2O as well as CO_2 and SO_2 are responsible for the comparatively high albedo. Furthermore, a group of materials termed tholins, consisting of, e.g., C–H and C–N-bearing constituents are assumed to be abundant on both satellites, creating dark-red coloured material on these surfaces.

On global and regional scale, Callisto is dominated by densely cratered plains with little geologic modification, as shown in Fig. 12.1a. The high crater density indicates a high surface age on the order of >4 Ga in existing cratering chronology

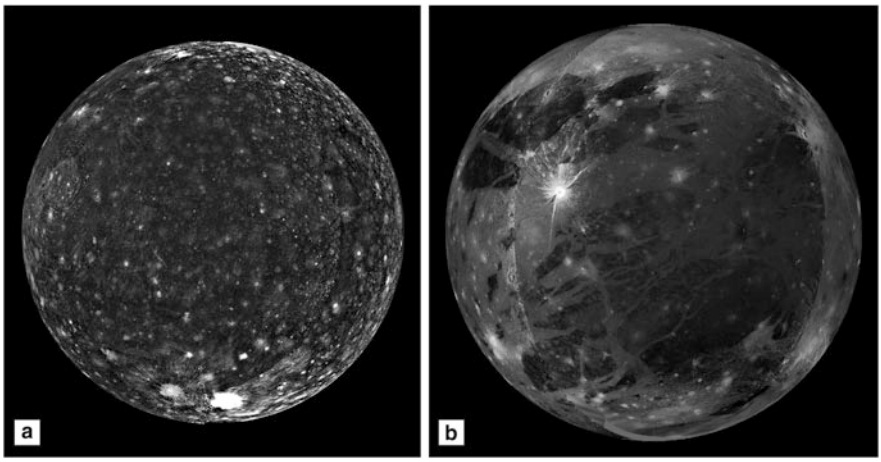

Fig. 12.1 The two outermost and largest Galilean satellites of Jupiter. (**a**) Callisto and (**b**) Ganymede, shown with their Jupiter-facing hemispheres in exact size ratio. Source: (**a** and **b**) NASA/Galileo SSI Team/DLR

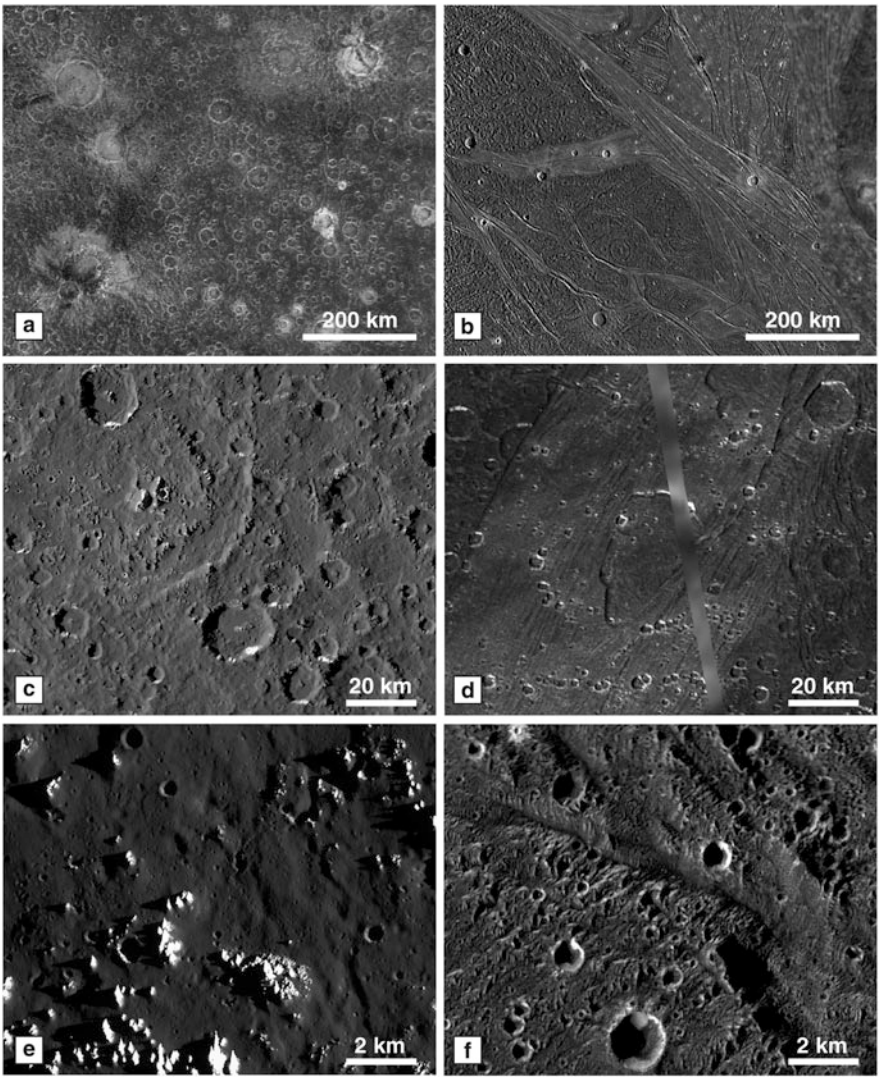

Fig. 12.2 Details of the major geologic units on Callisto (*left*) and Ganymede (*right*) at three different spatial resolutions. (**a**) Callisto at low resolution (36°S, 74°W, 950 m/px); (**b**) Ganymede at low resolution (39°N, 190°W, 950 m/px); (**c**) Callisto at intermediate resolution (8°N, 6.3°W, 160 m/px); (**d**) Ganymede at intermediate resolution (24°S, 318°W, 160 m/px); (**e**) Callisto at high resolution (0.85°N, 106.2°W, 15 m/px); (**f**) Ganymede at high resolution (16°S, 309°W, 15 m/px). Source: (**a**–**f**) NASA/Galileo SSI Team/DLR

models. The cratered plains appear mostly dark in Voyager and Galileo SSI images but are superimposed with numerous craters in a wide range of albedos, preservation states and ages. Bright craters, in some cases with rayed ejecta, are the youngest impact features. The majority of craters is dark and degraded which provokes the

generally low albedo of the cratered plains. In addition to craters, large multi-ring basins with total diameters of several thousand kilometers are abundant.

Ganymede (Fig. 12.1b) displays two major surface units distinguishable by albedo and morphology. About one third of the surface is characterized by dark, densely cratered plains similar to the global dark plains on Callisto, also comparable in cratering model age. Systems of arcuate furrows represent the remnants of former multi-ring basins similar to those seen on Callisto but the basins on Ganymede were resurfaced by younger geologic activity. The second major geologic unit on Ganymede which covers about two thirds of its surface are bright plains which are less densely cratered, younger and heavily tectonized. These bright plains consist of longitudinal swaths termed Sulci and polygons several tens of kilometers across. Most of the bright terrain shows a tectonic imprint of numerous closely spaced parallel grooves. Only a small amount of bright polygons appears smooth at Voyager resolution.

In Fig. 12.2, the surfaces of Callisto and Ganymede are compared at three different levels of spatial resolution. At 950 m/px, differences even in the dark cratered plains on both Callisto and Ganymede become apparent. Callisto (Fig. 12.2a) is more densely cratered than Ganymede but cratering model ages are more or less the same since the cratering rate is higher at Ganymede due to gravitational focusing by Jupiter. A wide range of impact-crater morphologies not known from terrestrial planets can be distinguished. At this resolution, segments of furrows are revealed indicating remnants of heavily degraded multi-ring basins. At same scales, the dark cratered plains on Ganymede differ remarkably from those on Callisto by a higher degree of resurfacing due to degradation and/or tectonism (Fig. 12.2b). Craters appear more degraded than on Callisto, preferentially by furrows tectonic and/or impact-tectonic in origin. The bright plains consist of linear swaths, either grooved or smooth. Since the Voyager flybys bright terrain was known to have formed at the expense of dark terrain which is documented in bright groove lanes transecting and even dissolving dark terrain.

At higher resolution of ∼160 m/px, Galileo SSI images show that a large number of craters in the dark plains on Callisto are heavily degraded with rims having been dissected into groups of blocks or massifs (Fig. 12.2c). Only a small number of craters appears still sharp. Furthermore, the surface is covered by a blanket of dark material which is ubiquitous on Callisto. Galileo SSI data at higher resolution, here ∼160 m/px in Fig. 12.2d have shown that the Voyager-based notion of bright terrain formation by tectonic extension, cryovolcanic flooding and subsequent deformation creating numerous grooves had to be replaced by a formation process which is predominantly tectonic. Dark terrain is transformed into bright terrain by a process termed tectonic resurfacing. Cryovolcanism is thought to play a minor role instead. However, landforms reminiscent of volcanic calderas are found in a small number of localities, like the one in the center of Fig. 12.2d. The crater density on the bright terrain is in fact lower than that in the dark plains. The misleadingly large number of craters in chains and clusters trending approximately west–east as shown in Fig. 12.2d stems from secondaries whose source crater is located further to the west.

High-resolution Galileo images taken at 15 m/px reveal the dominance of erosion and degradation of landforms such as crater rims on Callisto, concomitant with small impacts (Fig. 12.2e). This degradation is thought to be caused by diurnal variations in solar insolation, causing the sublimation of volatiles and leaving a residual which forms a globally abundant dark smooth lag or deposit. Bright, ice-rich material remains in the form of massifs or groups of massifs. Unlike Callisto, Ganymede at 15 m/px resolution appears to be less affected by degradation than does Callisto, especially the bright regions (Fig. 12.2f). The area depicted in detail covers a ∼NW–SE-trending boundary between a bright grooved polygon (E) and a bright region smooth at Voyager resolution (W). Even in the presumed smooth terrain, fine-scaled tectonism is visible, as in the grooved terrain.

Specific morphology of impact craters is a feature which the two otherwise dissimilar icy moons Callisto and Ganymede have in common. Both satellites show the widest range of impact crater morphologies compared to other planets or icy satellites. Some of these forms do not occur on the terrestrial planets therefore rheological properties of icy surfaces are different than those of the terrestrial planets.

On both Callisto and Ganymede, small bowl-shaped craters and complex craters with central peaks occur. The crater forms which are not abundant on terrestrial planets are shown in Fig. 12.3. Pedestal craters like *Achelous* are forms which are characterized by outward-facing scarps at the distal part of continuous ejecta. Dome craters like *Melkart* show a central dome within a rimmed central pit. Anomalous dome craters, otherwise termed large dome craters or pene-palimpsests such as *Neith* display a heavily dissected wreath-like broad rim surrounding a large dome while the nominal crater rim can barely be identified and is on average about twice the diameter of the dissected rim diameter. Another type of impact structure first seen in Voyager images comprises bright pancake-shaped features, preferentially in Ganymede's dark terrain, with little topographic expression of, e.g., a crater rim. These specific landforms were termed palimpsests or palimpsest craters. Higher-resolution Galileo SSI images reveal concentric structures reminescent of crater rims and remnants of central pits, as in *Buto Facula*. Dome craters and pene-palimpsests on Callisto can also be identified in Fig. 12.2a. The most complex impact structure forms are multi-ring basins on Callisto which were not altered by subsequent geologic processes. In Fig. 12.4 a detail of basin *Valhalla* is shown which can be subdivided into structural zones with numerous scarps, ridges and troughs not known from ring basins on terrestrial planets.

The formation of these specific crater forms and the origin of tectonic features on Ganymede are still poorly understood. Prerequisite for these landforms to occur is the stratification into a brittle ice shell overlying a layer of more mobile, viscous ice. The thickness of these two strata varies from body to body. Galileo magnetometer measurements revealed induced magnetic fields, indicating the existence of water oceans at depth. For Ganymede the induced field strength is exceeded by the strength of the magnetic field created in a dynamo in its core which makes the satellite the only one to have a self-generated magnetic field. The oceans could be at a depth of ∼100–300 km at Callisto, and ∼170 km at Ganymede and are estimated to be

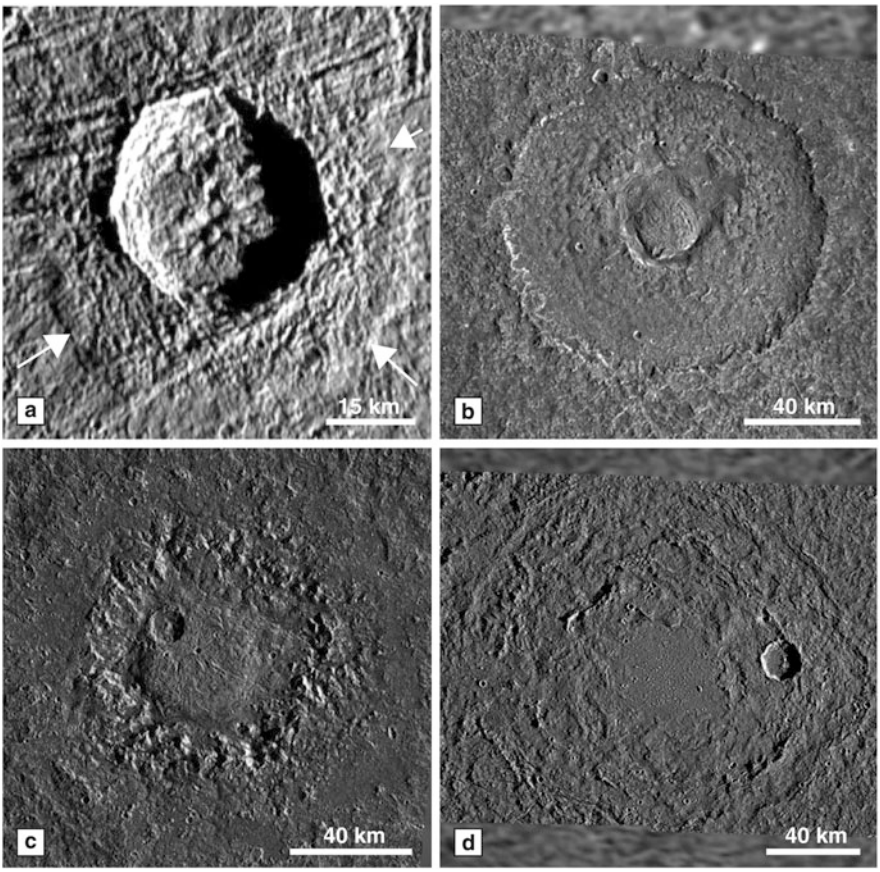

Fig. 12.3 Four types of craters which occur on Ganymede (as shown) as well as Callisto but are not abundant on terrestrial planets; (**a**) pedestal crater *Achelous* (61.8°N, 11.7°W, ⊘35 km diameter); (**b**) dome crater *Melkart* (9.9°S, 186.2°W, ⊘105 km); (**c**) anomalous dome crater (or pene-palimpsest) *Neith* (29.4°N, 7.0°W, ⊘135–140 km); (**d**) palimpsest *Buto Facula* (13.2°N, 203.5°W, nominal crater rim diameter 245 km). Source: (**a**–**d**) NASA/Galileo SSI Team/DLR

several tens or hundreds of kilometers deep. The origin and age of Ganymede's bright grooved terrain is an open question, taking into account the uncertainties of cratering model chronologies. Currently, Ganymede is in an orbital resonance with Io and Europa (Laplace resonance) but tidal stress is not sufficient to create tectonic deformation at present. A chaotic orbital evolution prior to the Laplace resonance in the past and/or differentiation processes involving phase changes of ice in the interior might have created predominantly extensional stress to have been responsible for global tectonic resurfacing. The grooved terrain could be ∼2 Gyr years or even up to ∼4 Gyr years old, depending on the cratering model chronology used while the dark terrain is estimated to be >4 Gyr.

Fig. 12.4 Detail of *Valhalla*, the largest multi-ring impact basin on Callisto. Impact structures like *Valhalla* are characterized by a central bright plains unit, surrounded by numerous concentric rings of inward-facing scarps and troughs. Source: NASA/Galileo SSI Team/DLR

12.1.2 Europa: A Heavily Tectonized Ice–Rock Satellite

In the class of icy satellites Europa is a special case because of its comparably high mean density of 3013 kg/m^3. Europa has a diameter of 3124 km and is about the size of the moon. The high mean density implies that it has a higher content of heavier constituents, e.g., silicates, and therefore is classified as an ice-rock satellite. Europa's surface is bright with an average albedo of ∼0.64. The dominant spectral species on its surface detected in the near-infrared is water ice. Further species spectroscopically detected with the Galileo NIMS mapping spectrometer are SO_2, CO_2, O_2, H_2O_2, tholins, and significant amounts of hydrated salts.

The smallest of the four Galilean satellites has by far the most complex surface morphology. Global colour views (Fig. 12.5) show the two main geologic units: bright bluish plains and brown mottled plains. Tectonic landforms, mainly long dark-brown lineaments dominate the surface.

At higher resolution, the bright, in colour bright-bluish, plains are seen to consist of numerous parallel ridges in lanes or polygons, much alike the grooved terrain on Ganymede. Similarly, ridges are seen even at the highest resolutions (Fig. 12.6). Prominent ridges several hundreds or even thousands of kilometers long as those shown in the global views in Fig. 12.5 represent a specific feature termed *double ridges*, characterized by a medial trough flanked by a pair of two ridges. From stereo analysis of Galileo SSI images heights of several hundred meters could be derived for one of the prominent double ridges). Several models of double ridge formation exists, involving tectonism, cryovolcanism, or a combination of the two processes.

Another prominent landform associated with bright ridged plains are dark, in colour dark brown, wedge-shaped bands as the one shown in Fig. 12.6.

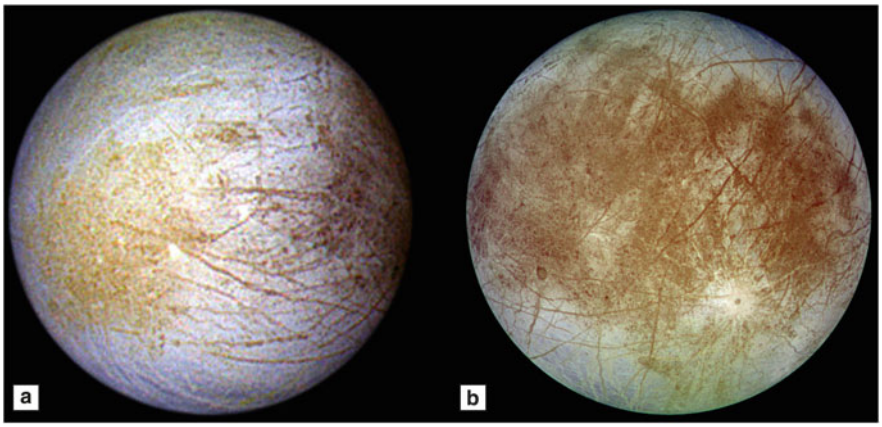

Fig. 12.5 Global color images of Europa, showing mainly (**a**) the leading hemisphere (PIA01295) and (**b**) the trailing hemisphere (PIA00502), both approximately in natural color. Source: (**a**) (PIA 01295): NASA/JPL/University of Arizona. (**b**) (PIA 00502): NASA/JPL/DLR

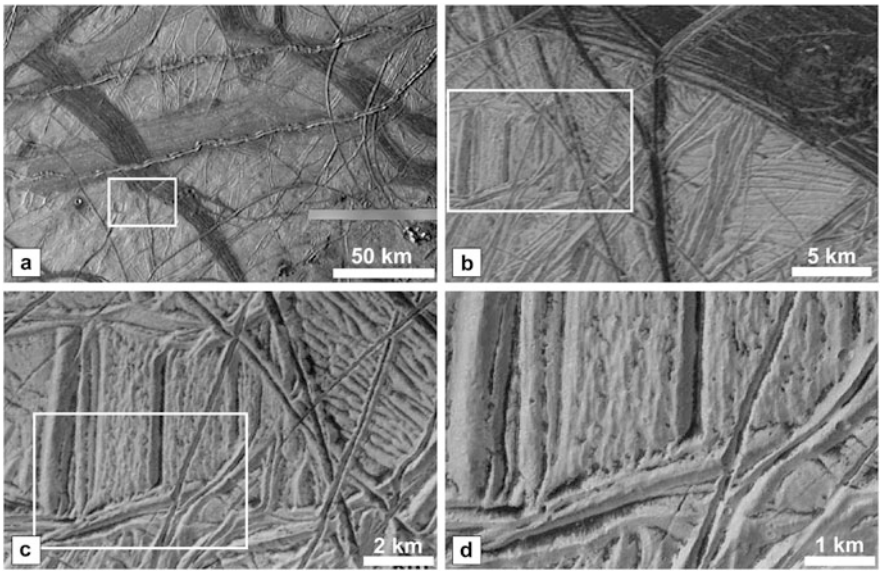

Fig. 12.6 Details of bright ridged planes and dark wedges on Europa, shown with increasing spatial resolution (location of each panel indicated by *rectangles*); (**a**) spatial resolution 430 m/px; (**b**) 55 m/px; (**c**) 25 m/px; (**d**) 12 m/px. Source: (**a**–**d**) NASA/Galileo SSI Team/DLR

Preferentially these features are concentrated on the anti-Jovian hemisphere. Stereo analysis has shown that wedge-shaped bands can be elevated ∼150 m above the surrounding terrain. At higher resolution, the dark or dark brown mottled terrain is resolved into a class of units specific on Europa termed *chaos regions*, covering

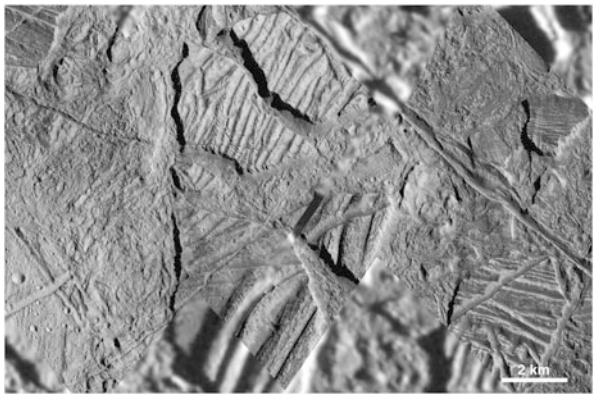

Fig. 12.7 Detail of *Conamara Chaos*, one example of chaos regions dominating the dark (*brown*) mottled plains. Plates of pre-existing terrain, mostly ridged plains, were translated and rotated within a hummocky matrix. Mosaic of Galileo SSI images at 10 m/px resolution in context of 55 m/px. Source: NASA/Galileo SSI Team/DLR

about 30% of the surface. This type of unit consists of disrupted remnants of pre-existing terrain, in the main plates or blocks of ridged plains which underwent translation and/or rotation in a hummocky matrix (Fig. 12.7). From stereo analysis, heights on the order of 200 m for these plates or blocks with respect to the matrix material have been derived. In general, the surface of Europa is very flat with height differences on the order of only several hundred meters. Topography does not exceed 1 km which is only achieved in a very small number of localities.

Europa has the least cratered surface of the three icy Galilean satellites. Depending on cratering model chronologies, the ages measured in specific units are on the order of only tens or hundreds of million years. Small impact craters resemble those on the Earth's moon but are flatter than these. The largest impact structures preserved on Europa are only 100–150 km across and show multi-ring morphology even at these small diameters.

The dominance of tectonic forms, the specific morphology and low topography of impact craters, and the low surface age are evidence that Europa may possess an internal water ocean whose existence has been inferred from an induced magnetic field measured by the Galileo magnetometer. This ocean is estimated to be on the order of at least $\sim$100 km thick and within a few tens or even kilometers from the surface. Most of the tectonic features can be explained by tidal stress acting on an ice shell overlying a liquid layer. In addition, patterns of tectonic forms indicate that Europa possibly rotates non-synchronously on the order of $\sim 10^5$ years which would require an ice shell decoupled from the interior through a liquid layer. Further tidal stress comes from the Laplace resonance between Io, Europa and Ganymede.

Since formation models of specific landforms such as, e.g., double ridges involve cryovolcanism and since the surface appears young it is possible that Europa is still active today. A search for candidate plumes during the Galileo mission was not successful, however. In December 2012, the Hubble Space Telescope was able to detect plumes of water ejected from the south polar region. Hence Europa most likely has to be classified as a cryovolcanically active icy satellite at present time.

12.1.3 Future Missions to the Icy Galilean Satellites

The global geologic mapping at high resolution and geophysical issues, such as, e.g., the existence of a subsurface ocean, are still incomplete or not solved. To achieve these goals, two future missions are currently being planned.

The *European Space Agency (ESA)* will send the *JUICE (JUpiter ICy moon Explorer)* spacecraft to Jupiter and Ganymede. Launch will be ∼2023, arrival and orbital phase is planned for 2030–2033. After two Europa flybys and several Callisto flybys the JUICE spacecraft will be inserted into a Ganymede orbit and be the first spacecraft to orbit a planetary satellite.

The *National Aeronautics and Space Administration (NASA)* is planning the *Europa Mission* with a launch scheduled for ∼2025 and an arrival after 2030. The spacecraft then will perform 45 flybys at Europa.

12.2 The Satellites of Saturn

Like Jupiter, Saturn has a large number of icy satellites; currently (mid-2015) a total of 62 is known of which 53 moons are named. Nine moons which are the largest ones were known prior to the flybys of the two Voyager spacecraft. With increasing distance these moons are Mimas, Enceladus, Tethys, Dione, Rhea, Titan, Hyperion, Iapetus and Phoebe. Unlike the Galilean satellites of Jupiter, their state of geologic evolution, surface ages and present-day activity are not correlated with the distance to Saturn. Except Hyperion and Phoebe, these satellites are in synchronous orbits around Saturn, facing one hemisphere always towards the central planet. In this section, the geology of the seven largest bodies is reviewed in detail (see Chap. 12 for details on Hyperion and Phoebe).

12.2.1 Mimas and Iapetus: Old Cratered Surfaces

Of the seven largest moons, Mimas is the innermost (and smallest) and Iapetus the outermost satellite. Both bodies are mainly characterized by densely cratered surfaces inferring a high surface age on the order of ∼4 Gyr and older. Mimas is a triaxial ellipsoid with a mean diameter of only 396 km. Its average density is 1152 kg/m^3. The high geometric albedo of 0.6, the domination of the spectrum by water ice-absorption bands as well as the low mean density infer that Mimas is a mostly icy body with a very small amount of heavier constituents. A global view of Mimas (Fig. 12.8a) shows a more or less uniform, densely cratered plains unit whose most remarkable characteristic is the large crater Herschel with ∼140 km diameter and a depth of ∼10 km. Colour features such as, e.g., an elliptically-shaped equatorial bluish band on the leading hemisphere (facing the direction of Mimas'

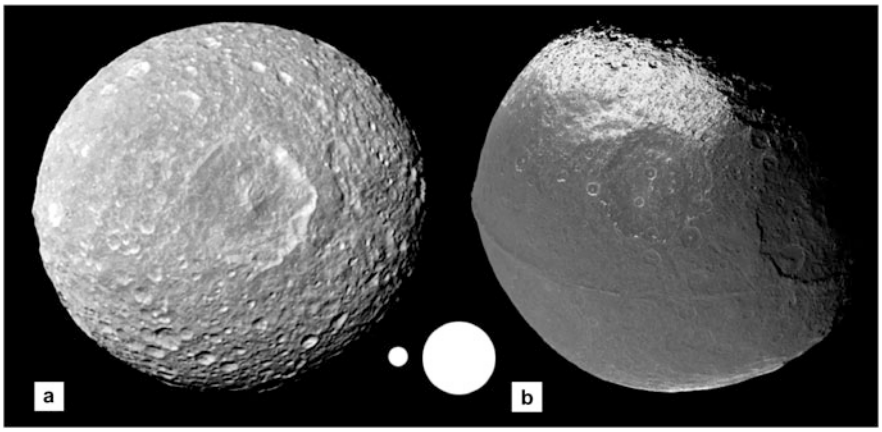

Fig. 12.8 Global views of (**a**) Mimas and (**b**) Iapetus; *filled circles* compare the respective sizes of the two moons. Largest crater on Mimas is Herschel on the leading side. Two of a number of impact basins, Falsaron and Turgis (near the terminator), are visible in the Iapetus mosaic showing mainly the dark leading hemisphere and the bright polar areas. The mosaic also shows Iapetus' remarkable equatorial ridge. Source: (**a**) (PIA 12568) NASA/JPL/Space Science Institute. (**b**) (PIA 06166) NASA/JPL/Space Science Institute

orbital motion) infer sputtering of particles from the Saturnian magnetosphere upon surface materials. Mimas has experienced little geologic evolution other than impact cratering and magnetospheric sputtering since the time of its formation.

Since its discovery by G.D. CASSINI in 1671, Iapetus has been known for its enigmatic albedo dichotomy between the leading and trailing hemisphere during its orbital period. On the leading side, Iapetus is extremely dark (geometric albedo: 0.05) while the trailing side and the polar regions are bright (albedo 0.5–0.6).

Iapetus is the third-largest moon of Saturn (1471 km mean diameter) with an average density of 1088 kg/m^3, hence close to that of water ice. Cassini ISS images have revealed that both dark and bright terrain are heavily cratered and include a number of impact basins several hundreds of kilometers across (Fig. 12.8b), more than on any other satellite of Saturn. Several impact basins could only be identified with the help of stereo imagery. Stereo data also demonstrated that Iapetus displays substantial topography inferring a very thick lithosphere. On average, the surface age of Iapetus is estimated to be the highest of the icy satellites ($>\sim$4 Gyr).

The most enigmatic topographic feature is an equatorial ridge which girdles almost the entire satellite (Fig. 12.8b). In some areas, the ridge is replaced by equatorially aligned peaks or massifs. Its origin is uncertain. Tectonic stress possibly in connection with tidal effects active in the earliest history of the satellite is discussed as a viable scenario for ridge formation. Spectral absorption features in the near infrared measured by Cassini VIMS and the preferentially red colour of the dark terrain in the visible spectrum infer compounds such as CO_2, O-H, C-H (aliphatic and aromatic hydrocarbons), Fe-bearing species and tholins to occur on Iapetus.

The dark, spectrally reddish material is from a still unknown external source. Thermal segregation is believed to mainly cause water ice evaporating from the leading side but not from the trailing side or from the polar regions, supported by the deposition of dark material preferentially on the leading side.

12.2.2 Tethys, Dione and Rhea: Impact Cratering and Tectonism

The three satellites Tethys, Dione and Rhea, in orbit about Saturn with increasing distance, share the following common characteristics

- they are nearly the same size (diameters 1066 km, 1124 km and 1528 km, respectively);
- spectrally, they feature significant water ice abundances,
- they show vast regions of densely cratered plains, including the presence of large impact features (basins), and
- tectonic features occur, regionally and locally. Figure 12.9 shows the tectonically resurfaced parts of their surfaces.

Tethys has the lowest mean density of the major satellites, only 956 kg/m^3. Its spectral properties are dominated by deep water ice absorptions comparable to those on Enceladus. Globally, Tethys is heavily cratered, including several large craters (basins) of which the largest one is Odysseus (445 km diameter). The most remarkable feature on Tethys is the graben system of Ithaca Chasma which encircles almost the whole satellite (Fig. 12.9a). The graben roughly follows a great circle. Ithaca Chasma is a terraced trough, approximately 100 km wide with narrower

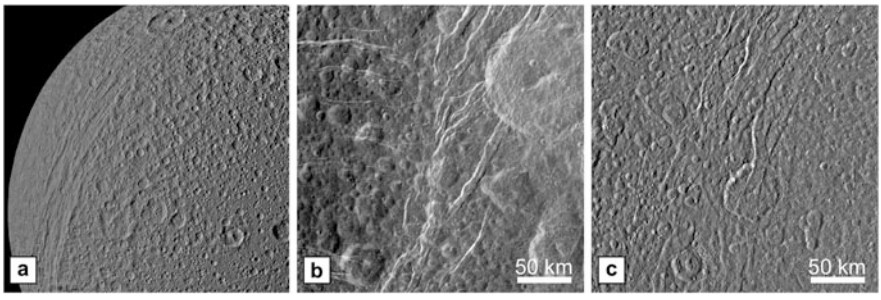

Fig. 12.9 Tectonic features on Tethys, Dione and Rhea; (**a**) densely cratered plains on Tethys and the graben of Ithaca Chasma. Largest crater (top of image) superimposing the tectonic structures is Telemachus (92 km diameter; 54°N; 339.4°W); (**b**) graben system Eurotas Chasmata (approximately west-east) on Dione, truncated by younger graben system Padua Chasmata (approximately north–south), superimposed by crater Ascanius (largest crater; 98 km diameter, 33.4°N; 232.2°W); (**c**) north–south–trending graben system Avaiki Chasmata on Rhea, cutting crater Kuma (largest crater; 50 km diameter, 10°N; 277.2°W)

branches. Stereo analysis shows it is ∼3 km deep and possesses raised flanks up to 6 km in height above the surrounding terrain. It has been suggested that the impact of Odysseus was responsible for the tectonic stress which created Ithaca Chasma. Crater counts on Tethys, however, infer that Ithaca Chasma is relatively older than Odysseus. When Ithaca Chasma was created is highly uncertain. Cratering model ages derived from crater counts are on the order of 3–4 Gyr.

Since a formation of Ithaca Chasma by impact-induced tectonism seems unlikely, other processes are responsible. Apart from early differentiation processes, paleo-resonances with other moons, e.g., Dione, could have caused a higher orbital eccentricity, creating tidal stress able for tectonic features to form.

The average density of Dione, 1469 kg/m^3 is higher than that of its two neighbours Tethys and Rhea, indicating a higher content of heavier material in its bulk composition. The leading hemisphere has a higher albedo and a higher water content than the trailing hemisphere. In turn, the trailing side has a higher amount of dark contaminants and an enhanced CO_2 absorption, most likely caused by bombardment of magnetospheric particles. Dione's surface is densely cratered, especially its leading hemisphere, and also includes large impact structures (basins). Contrarily, the trailing hemisphere is characterized by a concentration of tectonic graben systems termed chasmata with a variety of trends. Scarps of the graben and troughs in this hemisphere are bright and appear as a dense network of linear or slightly curved markings in images with lower resolution (a feature termed wispy terrain after the Voyager flybys). The high albedo is attributed to an exposure of water ice along the graben walls.

Apart from Enceladus, Dione displays the highest abundance of tectonic landforms of the major Saturnian satellites. The trends of tectonic graben and troughs, their mutual crosscutting and/or superposition infers varying stress systems with time. Figure 12.9b shows an example for two sets of graben, an older west–east trending one, and one younger trending north-south and truncating the older system. A sequence of tectonic events or episodes can be established by detailed geologic mapping.

As for the other satellites, absolute age assignments from crater counts are model-dependent and very uncertain. The densely cratered units may be as old as 4–3 Gyr while the tectonic resurfacing could have taken place between ∼3 and ∼1 Gyr ago. The origin of these tectonic features is still not clear. Episodic paleo-resonances involving higher eccentricities of Dione's orbit could have created the necessary tidal deformation to repeatedly fracture the surface with different trends.

Rhea is the second-largest moon of Saturn (1528 km diameter) and has an average density of 1233 kg/m^3. Like Dione, it has a pronounced leading-trailing asymmetry in albedo. Similarly, water ice is the dominating surface component, contaminated by dark material whose exact composition is in detail unknown. Dark material and a minor amount of CO_2 are concentrated on the trailing side caused by the impact of dust and magnetospheric particles.

Rhea features vast expanses of densely cratered plains with little geologic diversity, except for the trailing hemisphere. Large impact craters (basins) are abundant. Like on its inner neighbour Dione, systems of troughs and graben occur

on the trailing side (Fig. 12.9c). At lower resolution, these appear as bright linear markings (termed wispy terrain, as on Dione). Exposure of water ice along graben walls causes the high albedo in this tectonized area. Unlike Dione, however, the graben system on Rhea is more uniform with a major north–south trend, inferring there was only a single episode of tectonic deformation.

Crater counts and the application of cratering chronology models yield similar ages as for Dione's densely cratered plains and tectonic features. Youngest surface feature by means of crater counts is the bright ray crater *Inktomi* (47.2 km diameter, 14.1°S, 112.1°W) with an age on the order of hundreds of millions, or only millions of years, depending on the model chronology.

12.2.3 Enceladus: A Small Active Icy World

Enceladus, orbiting Saturn between Mimas and Tethys, has only a diameter of 504 km but a mean density of 1606 kg/m^3 indicating heavier material contributing to its bulk composition. Its geometric albedo is 1.0, higher than for any other satellite. Spectra of Enceladus are dominated by water ice with minor amounts of CO_2, organics, and possibly NH_3.

The surface is characterized by a wide range in landforms, unlike the other Saturnian airless moons, and includes (a) craters, (b) tectonic features, and (c) features indicative of cryovolcanism. Figure 12.10 shows densely and moderately

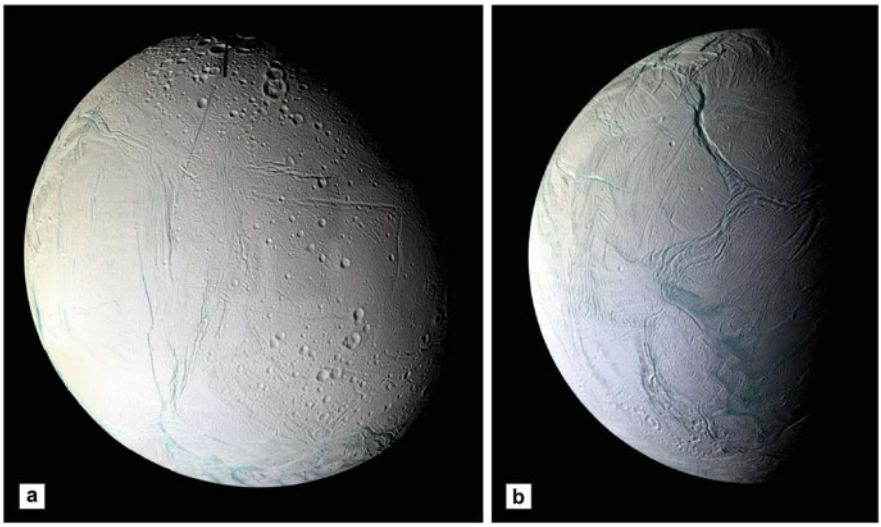

Fig. 12.10 Global colour Cassini ISS mosaics of Enceladus; (**a**) trailing hemisphere showing densely and moderately cratered areas cut by numerous linear or curved tectonic features (PIA08353); (**b**) south polar terrain; the bluish linear troughs in the lower part of the colour mosaic extending into the unilluminated part represent the source region of active cryovolcanism (PIA11133). Source: (**a**) (PIA 08353) NASA/JPL/Space Science Institute. (**b**) (PIA 11133) NASA/JPL/Space Science Institute

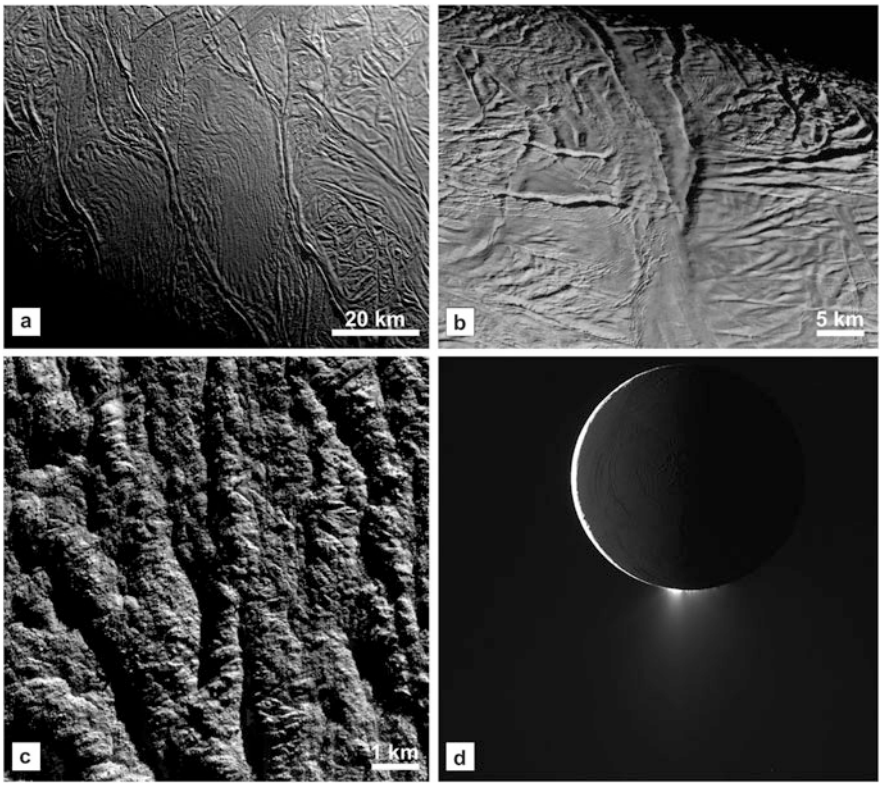

Fig. 12.11 Tectonic landforms on Enceladus, seen at various spatial resolutions; (**a**) troughs or grooves with raised rims in the south polar terrain (so-called tiger stripes), 100 m/px resolution; (**b**) oblique view of troughs and ridges in the south polar terrain, 45 m/px resolution; (**c**) low-sun image taken from the south polar terrain at 9 m/px resolution, revealing small-scale tectonism; (**d**) cryovolcanic plumes erupting from the tiger stripes in the south polar terrain

cratered plains, linear or curved tectonic graben or ridges, and smooth terrain in global colour mosaics. Terrain with higher crater frequencies is concentrated in the northern hemisphere and in the mid-latitudes. Tectonically resurfaced areas are generally low or even devoid in craters implying very young surface ages. Tectonic forms, troughs, grooves and ridges, are detectable at all spatial resolutions (Fig. 12.11a–c).

In 2005, active cryovolcanism in the south polar terrain was discovered. Plumes of water ice particles episodically erupt from linear groove-like features (termed *tiger stripes*), feeding the E-ring whose material is spread along the orbital path of Enceladus (Fig. 12.11a, d).

The cryovolcanism is mainly driven by intense tidal forces. Liquid reservoirs of water, possibly mixed with ammonia, are assumed to exist at depths of 500 m to

2 km beneath the surface. Stereo image analysis in other tectonized areas infer a thickness of ∼2 km of the brittle lithosphere.

The cryovolcanically active region is associated with a high heat flow across the tiger stripes as measured by instruments aboard Cassini. Analyses using Cassini VIMS data revealed a correlation between surface age and water ice particle sizes which are largest at the center of cryovolcanic activity. Smallest particle sizes are found in the old cratered terrains, tectonically resurfaced regions older than the south polar terrain show intermediate particle sizes. Exposure towards space weathering processes including bombardment with magnetospheric particles is a viable explanation for changes in particle sizes with time.

Enceladus is the Saturnian moon with the widest range of surface ages obtained by crater counts, from densely cratered plains with ages on the order of several 10^9 years to the south polar terrain younger than 10^6 years, depending on the chronology model applied.

12.2.4 Titan: A Large Earth-Like Satellite

Saturn's largest satellite Titan has a diameter of 5150 km and, together with Ganymede and Callisto, forms a trio of the largest and planet-sized moons in the Solar System. Its average density of 1880 kg/m^3 implying a rick-ice mixture is also similar to the densities of the two largest Galilean satellites. Titan is a unique satellite since its surface is shrouded behind a dense atmosphere which is opaque at visible wavelengths. The atmosphere can only be penetrated with optical sensors in windows in the near infrared as with the Cassini VIMS mapping spectrometer and with the Cassini synthetic aperture radar (SAR) instrument. The atmosphere is primarily composed of N_2 with some amount of CH_4.

On global scale, two major terrain types identifiable in near-infrared images occur on Titan: (a) bright terrain and (b) dark terrain. In general, the bright and dark terrains correlate with regions which also appear bright and dark at radar wavelengths. Dark terrain is concentrated in the equatorial regions while bright terrain is abundant in the mid- and high latitudes. In terms of geologic diversity, Titan is remarkably similar to Earth and is characterized by processes not active on airless icy satellites. Exogenic processes on Titan comprise (a) impact cratering, (b) erosion, transport and deposition by wind (eolian processes), and (c) erosion, transport and deposition by the presence of liquid material (fluvial and lacustrine processes). Endogenic processes on Titan include tectonism and cryovolcanism.

The number of impact craters identified in the surface area covered by the Cassini Radar instrument is on the order of ∼50 with only ∼8 craters believed to be *true* impact craters, such as crater *Momoy* (Fig. 12.12a). Features confirmed as true impact craters are several tens up to hundreds of kilometers in diameter. The low number of impact craters implies a comparable young surface age of hundred million years. Absolute surface ages on Titan, however, are an open question. The presence of liquid hydrocarbons (methane, ethane), thought to occur on Titan

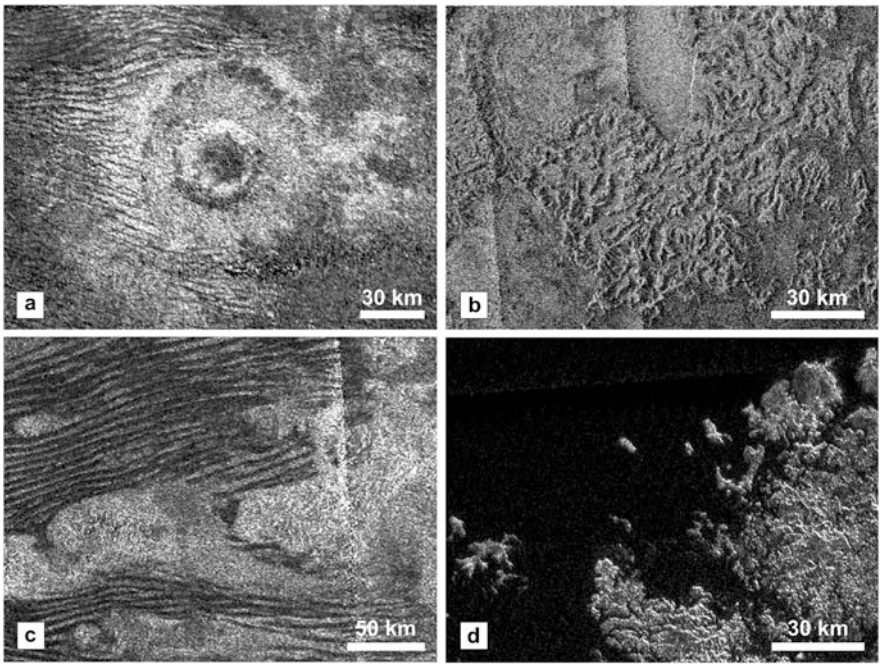

Fig. 12.12 Titan surface features I, imaged by the SAR radar instrument aboard Cassini. Details extracted from images in planetary photojournal; (**a**) impact crater Momoy (⊘40 km, 11.6°N; 44.6°W, PIA14744); (**b**) landscape shaped by erosion (PIA10219); (**c**) dunes (PIA09181); (**d**) Ligeia Mare, lake filled with liquid carbohydrates (PIA09211). Source: (**a**) (PIA 14744) NASA/JPL-Caltech/ASI. (**b**) (PIA 10219) NASA/JPL-Caltech/ASI. (**c**) (PIA 09181) NASA/JPL-Caltech/ASI. (**d**) (PIA 09211) NASA/JPL-Caltech/ASI

prior to Cassini, were confirmed by VIMS and the SAR radar instrument. Also, a hydrological cycle exists on Titan, with liquid carbohydrates playing the role of water on Earth. Subsurface liquids as well as wind on Titan produce etched terrain, landforms reminescent of terrestrial karst, as shown in Fig. 12.12b.

In the dark, topographically low-lying equatorial regions, longitudinal dunes are an ubiquitous landform. The majority of these dunes is west–east-oriented, suggestive of a dominant wind direction (Fig. 12.12c). Dunes extend for several tens or hundreds of kilometers, are on the order of 30–70 m high and diverge at, or are blocked by, topographic obstacles, as the bright massifs seen in Fig. 12.12c.

Dark areas smooth at radar wavelengths suggest standing bodies, or lakes, of liquid material, preferentially in the northern and southern high latitudes. A specular reflection observed for the first time in 2009 in VIMS data was shown to originate from one of the presumed lakes in the northern regions and unequivocally confirmed the existence of liquid material. A number of lakes contain liquids, such as *Ligeia Mare* (Fig. 12.12d) but dry lake beds also occur. These lakes in parts cover huge areas comparable to, e.g., the Great Lakes in northern USA.

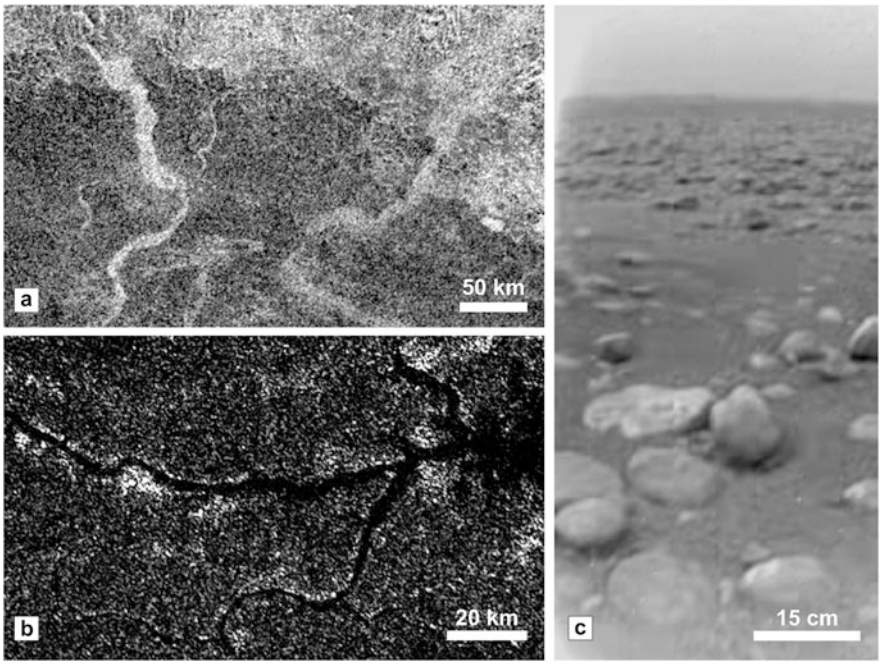

Fig. 12.13 Titan surface features II, imaged by the SAR radar instrument aboard Cassini. (**a**) Dry rivers in Xanadu region (PIA10956); (**b**) river flowing into Ligeia Mare (PIA16197); (**c**) Huygens landing site taken by the DISR instrument aboard the Huygens probe. Source: (**a**) (PIA 10956) NASA/JPL-Caltech/ASI. (**b**) (PIA 16197) NASA/JPL-Caltech/ASI. (**c**) NASA/JPL/ESA/University of Arizona

Valley systems carved by rivers of liquid carbohydrates can be observed in Cassini radar, VIMS, ISS (near-infrared filters) data as well as in images of the DISR instrument aboard the Huygens Titan landing probe. These valleys are formed by precipitation and runoff of rainfall, most likely of liquid methane. Different morphologic types, including dendritic systems can be identified. Dry valleys similar to terrestrial desert washes were created by episodic precipitation (Fig. 12.13a). River beds which appear dark represent active rivers such as the dendritic system emanating into *Ligeia Mare*, shown in Fig. 12.13b).

ESA's Huygens probe landed in a dry lake bed on January 14, 2005. The surface around the lander (Fig. 12.13c) reveals rounded blocks of ice, indicative of transport and deposition by a liquid material as well as fine-grained darker material (sand) transported and deposited by wind. The existence of endogenic activity on Titan, either past or even at present time is not unequivocally accepted and an issue in discussion. Mountains, circular domes and flow-like features indicate possible tectonic and cryovolcanic activity. Linear ranges of mountains could have been formed by compressive and/or extensional tectonism. Alternatively, Titan's surface could be shaped exclusively by exogenic processes.

12.2.5 Impact Crater Forms on the Saturnian Satellites

In general, impact craters on the Saturnian satellites differ from those on the Galilean satellites Ganymede and Callisto. Most of the Saturnian moon craters have central peaks while central pits or domes are absent. Also, no palimpsests are abundant. Broadly, craters on the Saturnian satellites resemble those on the terrestrial planets but are in general flatter indicative for a weaker icy target material compared to silicates. The surfaces of the Saturnian moons are also devoid of large multi-ring impact structures (such as, e.g., Valhalla on Callisto). In some large impact craters of basin size (>150–200 km in diameter), an inner ring may be present, however. The more terrestrial planet-like morphology of craters on the moons of Saturn is primarily attributed to the lower surface temperatures and, hence, higher strength of surface materials compared to the rheologic properties on Ganymede or Callisto.

12.3 The Satellites of Uranus

The five major icy satellites of Uranus, with increasing distance from the planet Miranda, Ariel, Umbriel, Titania and Oberon, were imaged during the Voyager-2 flyby on January 24, 1986. The long exposure duration of Voyager images necessary at the large distance of Uranus from the Sun and flyby distances to the targeted moons reduced the number of images suitable for detailed geologic analysis considerably, except for Miranda and Ariel. Only the southern hemispheres of all five satellites could be imaged while their northern regions remain more or less unknown terrain. Due to the rotational axis of Uranus (including the orbits of its major moons) being tilted by about 90°, the terminator lies near and approximately parallel to the equator. Until today, a total of 27 moons are known, most of them icy bodies much smaller than 200 km in diameter.

12.3.1 Oberon, Titania and Umbriel

Titania and Oberon are the two largest Uranian moons. Oberon (Fig. 12.14a) has a diameter of 1523 km and an average density of 1630 kg/m^3. Its surface is dominated by numerous impact craters, the largest of them often with central peaks, and with either bright or dark floors. In parts bright rayed ejecta occur representing young surface features. The dark floors were interpreted as candidate cryovolcanic deposits by the Voyager Imaging Team. Linear features reminescent of scarps and chasms several hundreds of kilometers long indicate tectonism, possibly extensional, but higher-resolution images are needed for detailed investigations.

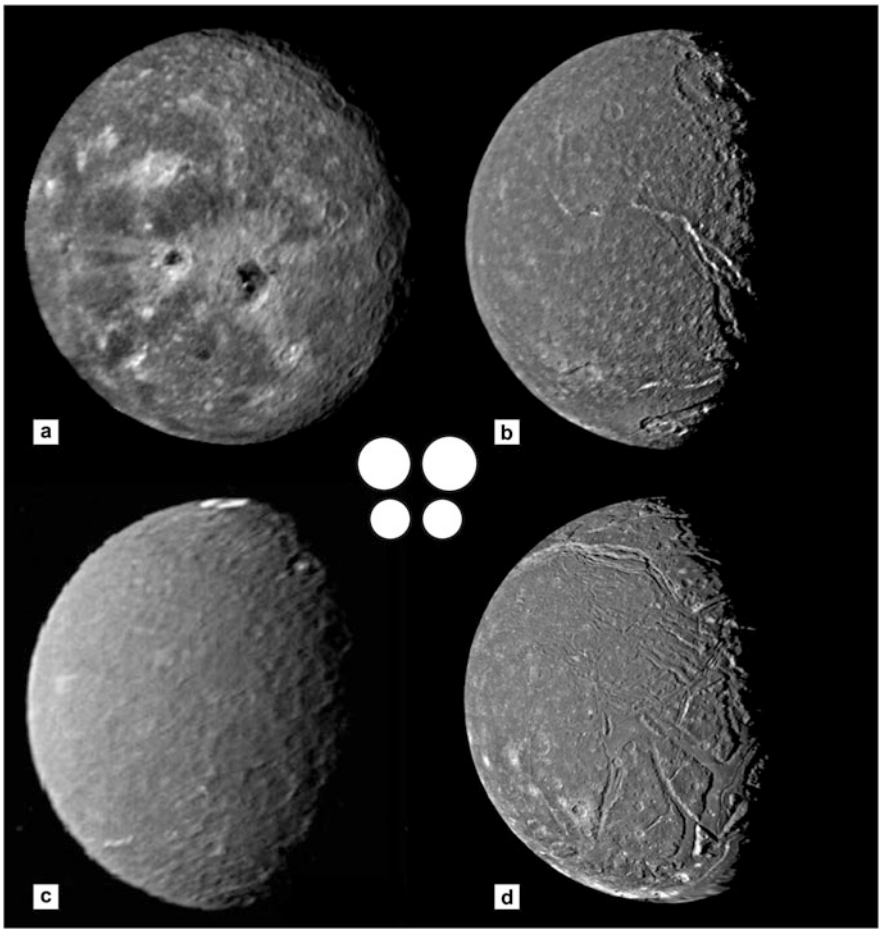

Fig. 12.14 The major Uranian satellites (**a**) Oberon (NASA/DLR), (**b**) Titania (PIA00039), (**c**) Umbriel (PIA00040), and (**d**) Ariel (NASA/DLR). These satellites are mostly densely cratered and therefore old, on the order of 4 Ga. Source: (**a**, **d**) NASA/DLR. (**b**) (PIA 00039) NASA/JPL. (**c**) (PIA 00040) NASA/JPL

Titania (Fig. 12.14b) measures 1578 km in diameter and has a mean density of 1744 kg/m^3. Like Oberon, Titania is mostly densely cratered but appears to have experienced tectonism (predominantly extensional) more intensely than its outer neighbour. Also alike, many large craters have central peaks. The largest impact feature is Gertrude (Fig. 12.14b, located near the top at the terminator) which shows an inner ring rather than a central peak. Titania features a set of longitudinal scarps and graben as well as plateaus. Tectonic features appear much more pronounced than on Oberon. Smooth areas were thought to have a cryovolcanic origin but this interpretation is questionable and needs imaging data of much higher resolution.

Umbriel, 1170 km in diameter, mean density 1400 kg/m^3, is the darkest of the five major satellites. Its surface (Fig. 12.14c) is characterized by a more or less uniform albedo suggesting a possibly global blanketing or coating with dark material, either from a deposit of primordial material or, less likely, from cryovolcanic resurfacing. A bright ring seen at the uppermost limb (Fig. 12.14c) is related to a large impact crater named *Wunda*. Of the five major satellites, Umbriel seems to be the geologically least evolved of the five major moons of Uranus.

12.3.2 Ariel and Miranda

With a diameter of 1158 km, Ariel (mean density: 1665 kg/m^3) is almost the same size as its outer neighbour Umbriel but geologically completely different. Figure 12.14d shows a surface with impact craters and a significant abundance of tectonic landforms, indicative of intense tectonic resurfacing in the past. Sets of scarps, canyons, and graben, in parts with elevated rims as shown in a stereo anaglyph (Fig. 12.14d) crisscross the surface. The floors of the widest canyons are smooth and feature a medial trough, interpreted as an axial fissure from which viscous cryo-lava erupted on the canyon floor. Alternatively, cryovolcanic material could have been emplaced in lava tube-fed flows.

The small size of Miranda (mean diameter: 372 km; mean density: 1200 kg/m^3) is in strong contrast with the wide range of landforms abundant on its surface. Miranda was the best one imaged by Voyager 2 and features the highest state of geologic evolution of all five major Uranian satellites. Figure 12.15 shows densely cratered,

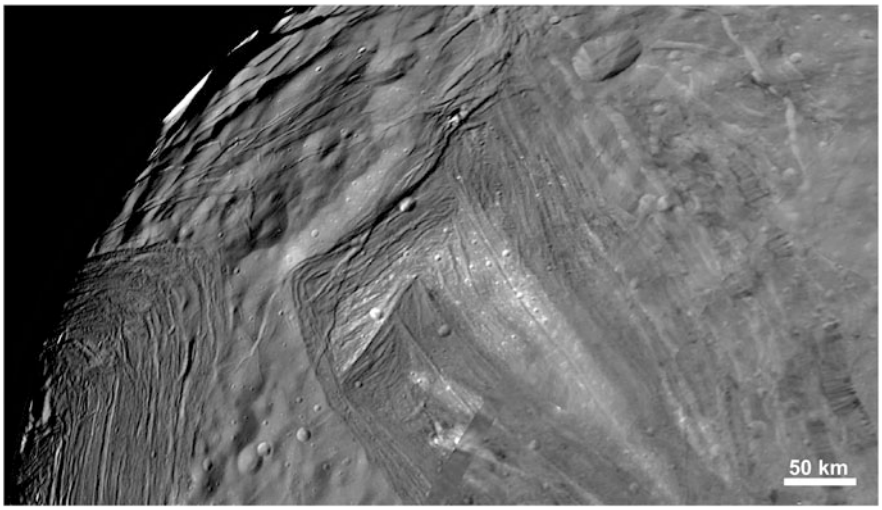

Fig. 12.15 Detail of a mosaic of Miranda showing cratered plains and tectonic landforms (Voyager 2)

little altered plains in association with sparsely cratered, highly modified terrain characterized by ovoidal- or trapezoid-shaped areas of parallel ridges and grooves, termed *coronae*. In addition, systems of parallel fractures and troughs and scarps up to 20 km in height are abundant.

12.4 Neptune's Largest Satellite Triton

Neptune and its icy satellites—in total 14 known today—were visited by the Voyager-2 spacecraft in its flyby in August 1989. Triton is unique in this satellite system: it is by far the largest moon, with a diameter of 2705 km even larger than the dwarf planet Pluto, and it orbits Neptune on a nearly circular orbit in about 6 days, but retrograde. Triton is believed to be a Kuiper Belt Object (KBO) captured by Neptune in the early Solar System. Another remarkable feature is the presence of a tenuous atmosphere with a surface pressure of $\sim 10^{-5}$ of the Earth's atmosphere, mainly of N_2 with CH_4 as a minor constituent.

During the Voyager-2 encounter only the southern hemisphere could be imaged in detail (highest resolution ~400 m/px). Spectroscopically, species such as N_2, CH_4, CO, CO_2 and H_2O were detected on the surface. Brown colours in Voyager colour images suggest the presence of tholins. Voyager-2 (Fig. 12.16) revealed a sparsely cratered surface, hence most of the surface covered by images is young, on the order of several tens of million years. The following geologic processes, past

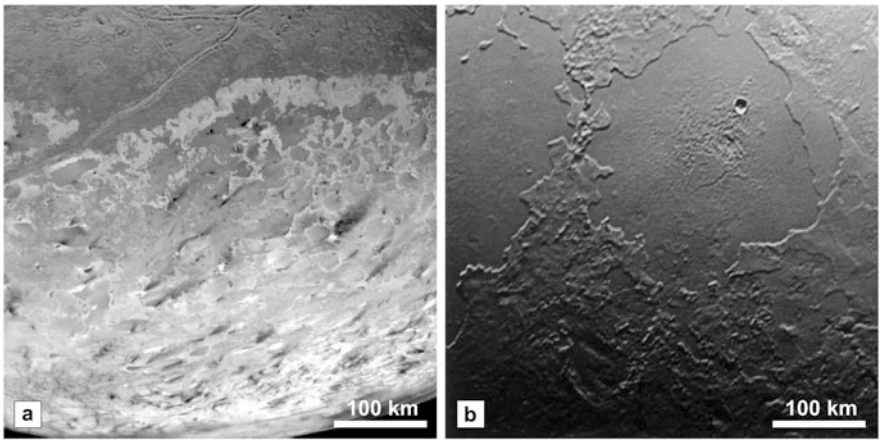

Fig. 12.16 Details of landforms on Neptune's largest satellite Triton; (**a**) tectonically altered terrain, characterized by double ridges (*top*), and a Triton-specific region termed *cantaloupe terrain*; (**b**) caldera-like landforms indicative of past cryovolcanism. Source: (**a**) (PIA 00059): NASA/JPL. (**b**) (PIA 01538): NASA/JPL

and present, characterize the surface: (a) erosion, (b) transportation and deposition by wind, (c) tectonism and (d) cryovolcanism. A Triton-specific terrain type termed *cantaloupe terrain* (Fig. 12.16a) consists of pits and dimples, possibly formed by either collaps, degradation and sublimation of surface materials, or, alternatively, by cryovolcanism. Dark streaks, seen in Fig. 12.16a extending from lower left to upper right (approximately SW–NE) represent wind-deposited material and infer a preferred wind direction. The prevalent tectonic landform are ridges of the double-ridge type also known from Europa, possibly with a similar origin (Fig. 12.16a on top of figure). Features resembling frozen lakes located within steep-sided calderas (Fig. 12.16b) could represent areas once filled with liquid material, similar to terrestrial lava lakes, and therefore are candidate sites of past cryovolcanism. Since Voyager-2 images of the limb of Triton revealed dark, vertically rising stems indicating possible geyser-like centers of activity the satellite might be classified as a volcanically active body at present.

12.5 Charon: Largest Satellite of Dwarf Planet Pluto

On July 14/15, 2015, the first spacecraft ever, New Horizons, flew by the dwarf planet Pluto (Chap. 13) and its five known satellites. Charon is the largest of Pluto's moons. Pluto and Charon can be considered a double planet circling in 6 days about a common center of mass with is located outside Pluto (unlike the center of mass of the Earth-Moon system). Both bodies exert considerable tidal stress to one another. Charon has a diameter of 1208 km, slightly more than half the size as Pluto, and a mean density of 1650 kg/m^3, similar to icy moons such as, e.g., Enceladus or Dione.

A global colour image of Charon (Fig. 12.17) shows a body with a darker polar area which has a sharp boundary towards its surrounding, and some albedo or colour variations in a generally bright, in colour grey terrain at the mid- and equatorial latitudes. Large craters are identifiable, especially such with bright rays and/or dark floors. A remarkable feature is an extended graben system spanning across the imaged hemisphere.

A higher-resolution image of the surface reveals impact craters, smaller-scale tectonism, and mountains, the largest of them surrounded by a moat-like feature. This specific morphology may indicate that the surface has strength not able to support larger loads. In summary, Charon is characterized by a wide range in surface features indicative of an intense geologic evolution.

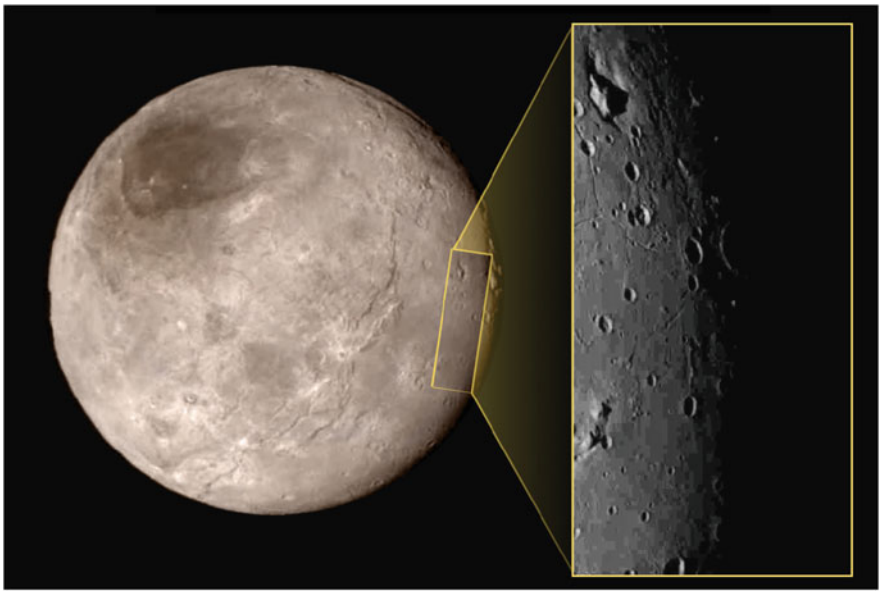

Fig. 12.17 Global colour image of Pluto's largest satellite Charon and a blow-up at higher resolution, obtained during the New Horizons flyby on July 14, 2015 by the LORRI camera. Source: (PIA 19713): NASA/Johns Hopkins University Applied Physics Laboratory/Southwest Research Institute

Take-Home Messages

Small, mid-sized and planet-sized moons orbit the four giant planets Jupiter, Saturn, Uranus and Neptune, grouped into a specific class of planetary objects termed *icy satellites* because of (a) the presence of water ice on their surfaces, spectroscopically detectable at near-infrared wavelengths, and because of (b) their average densities of less than $\sim$2000 kg/m^3.

Like the surfaces of the terrestrial planets (Chap. 9), their surfaces are shaped by three major geologic processes: (1) impact cratering (Chap. 7), (2) exogenic processes (Chap. 9), including, e.g., space weathering, erosion and degradation, and (3) endogenic processes (Chap. 8), including tectonism and cryovolcanism.

Jupiter's moon Callisto's surface is dominated by impact craters and degradation indicating an old surface (>4 Ga). There is a wide range of impact crater morphologies with little geologic modification except for degradation.

Jupiter's moon Ganymede shows a dichotomy with older dark cratered plains and younger bright tectonically resurfaced plains which was formed by tectonic resurfacing of dark terrain. Age of bright grooved terrain unknown, possibly 2 Ga and higher. Specific crater forms: dome craters, palimpsests, multi-ring basins.

Jupiter's moon Europa's surface is dominated by tectonic structures as well as landforms indicative of cryovolcanism. The density of impact craters is low implying a young surface age. So-called double ridges extending over 1000 km

are primary tectonic structures. About 30% of the surface are chaos regions, showing translation and rotation of plates of pre-existing terrain. Landforms indicate the presence of a subsurface ocean. Europa may be cryovolcanically active at present.

All airless Saturnian satellites feature old, densely cratered plains, either globally, or at least regionally (Enceladus)

The albedo dichotomy on Saturn's moon Iapetus stems from a combination of dark material deposition with thermal segregation of water ice preferentially on the leading side.

Tectonically resurfaced regions on Saturn's moons Tethys, Dione and Rhea possibly were caused by tidal stress during episodic orbital resonances with neighbour moons, involving higher forced eccentricities. Cryovolcanism and tectonism on Enceladus are driven mainly by strong tidal forces; material erupting from the south polar terrain feeds Saturn's E ring.

Liquid carbohydrates on Saturn's moon Titan form a hydrological cycle, like water on Earth. In addition to the geologic processes on airless icy satellites, Titan's surface is shaped by eolian, fluvial and lacustrine processes. Low number of impact craters indicate a low surface age of Titan (order of $\sim 10^8$ years).

The five major satellites of Uranus differ widely in geologic evolution. Oberon and especially Umbriel are geologically little evolved, while Miranda, Ariel and, to a minor degree, Titania show tectonically modified terrain. Tectonism is believed to be predominantly extensional, possibly due to volume changes in the thermal evolution of the satellites. Miranda is a special case and exhibits ovoids and the trapezoid which was possibly formed by extensional tectonism over upwelling plumes. An absolute time-scale when tectonism on these moons was active is not known or highly uncertain.

The surface of Neptune's moon Triton is sparsely cratered and therefore comparably young (order of several tens or hundreds of million years). Erosion and degradation could have formed the Triton-specific cantaloupe terrain. Double ridges on Triton possibly have the same origin as those on Europa. Dark deposits indicate a preferred wind direction in the tenuous atmosphere. Cryovolcanism was active in the past (calderas) and possibly is still going on today (geyser-like activity seen at the limb).

The range of surface features seen in the first images returned from Pluto's moon Charon imply an intense geologic history. Charon's surface features comprise impact craters with bright rays and/or dark floors, a large hemisphere-spanning graben system and mountains.

Suggested Readings

Brown, R., Lebreton, J.P., Waite, J.: Titan from Cassini-Huygens, viii + 535 pp. Springer, Dordrecht (2010). doi:10.1007/978-1-4020-9215-2

Carr, M.H., Belton, M.J.S., Bender, K., Breneman, H., Greeley, R., Head, J.W., Klaasen, K.P., McEwen, A.S., Moore, J.M., Murchie, S., Pappalardo, R.T., Plutchak, J., Sullivan, R., Thornhill, G., Veverka, J.: The Galileo Imaging Team plan for observing the satellites of Jupiter. J. Geophys. Res. **100**(E9), 18935–19955 (1995). doi: 10.1029/95JE00971

Croft, S.K., Soderblom, L.A.: Geology of the uranian satellites. In: Bergstrahl, T.H., Miner, E.D., Matthews, M.S. (eds.) Uranus, pp. 561–628. University of Arizona Press, Tucson (1991)

Dalton, J.B., Cruikshank, D.P., Stephan, K., McCord, T.B., Coustenis, A., Carlson, R.W., Coradini, A.: Chemical composition of icy satellite surfaces. In: Space Science Reviews, vol. 153(1), pp. 113–154 (2010). doi:10.1007/s11214-010-9665-8

Dougherty, M.K., Esposito, L.W., Krimigis, S.M. (eds.): Saturn from Cassini-Huygens. viii + 805 pp., Springer, Dordrecht (2009). doi:10.1007/978-1-4020-9217-6

Greeley, R.: Introduction to Planetary Geomorphology, 250 pp. Cambridge University Press, New York (2013)

Greeley, R., Chyba, C.F., Head, J.W., McCord, T.B., McKinnon, W.B., Pappalardo, R.T., Figueredo, P.: Geology of Europa. In: Bagenal, F., Dowling, T., McKinnon, W.K. (eds.), Jupiter – the Planet, Satellites, and Magnetosphere. Cambridge Planetary Science, vol. 1, pp. 327–362 (2004). Cambridge University Press, Cambridge

Jaumann, R., Clark, R.N., Nimmo, F., Hendrix, A.R., Buratti, B.J., Denk, T., Moore, J.M., Schenk, P.M., Ostro, S.J., Srama, R.: Icy satellites: geological evolution and surface properties. In: Dougherty, M.K., Esposito, L.W., Krimigis, S.M. (eds.) Saturn from Cassini-Huygens, pp. 637–681 (2009). Springer, New York

Lewis, J.S.: Physics and Chemistry of the Solar System, revised edition, 591 pp. Academic, San Diego (1995)

Melosh, H.J.: Planetary Surface Processes. Cambridge Planetary Science, vol. 13, 501 pp. Cambridge University Press, Cambridge (2012)

Moore, J.M., Chapmann, C.R., Bierhaus, E.B., Greeley, R., Chuang, F.C., Klemaszewski, J., Clark, R.N., Dalton, J.B., Hibbits, C.A., Schenk, P.M., Spencer, J.R., Wagner, R.: Callisto. In: Bagenal, F., Dowling, T., McKinnon, W.K. (eds.), Jupiter – the Planet, Satellites, and Magnetosphere. Cambridge Planetary Science, vol. 1, pp. 397–426. Cambridge Univ. Press, Cambridge (2004)

Müller-Wodarg, I., Griffith, C.A., Lellouch, E., Cravens, T.E. (eds.) Titan: Interior, Surface, Atmosphere, and Space Environment, 474 pp. Cambridge University Press, Cambridge (2014)

Pappalardo, R.T., Collins, G.C., Head, J.W., Helfenstein, P., McCord, T.B., Moore, J.M., Prockter, L.M., Schenk, P.M., Spencer, J.R.: Geology of Ganymede. In: Bagenal, F., Dowling, T., McKinnon, W.K. (eds.), Jupiter the Planet, Satellites, and Magnetosphere. Cambridge Planetary Science, vol. 1, pp. 363–396. Cambridge University Press, Cambridge (2004)

Prockter, L.M., Lopes, R.M.C., Giese, B., Jauman, R., Lorenz, R.D., Pappalardo, R.T., Patterson, G.W., Thomas, P.C., Turtle, E.P., Wagner, R.: Characteristics of icy surfaces. Space Sci. Rev. **153**(1), 63–111 (2008). doi:10.1007/s11214-010-9649-8

Schenk, P.M., Chapman, C.R., Zahnle, K., Moore, J.M.: Ages and interiors: The cratering record of the Galilean satellites. In: Bagenal, F., Dowling, T., McKinnon, W.K. (eds.) Jupiter – the Planet, Satellites, and Magnetosphere. Cambridge Planetary Science, vol. 1, pp. 427–456. Cambridge University Press, Cambridge (2004)

Stephan, K., Jaumann, R., Wagner, R.: Geology of icy bodies. In: Gutipati, M.S., Castillo-Rogez, J. (eds.), The Science of Solar System Ices, pp. 279–367. Springer, New York (2013)

Chapter 13
Small Bodies and Dwarf Planets

Nico Schmedemann, Matteo Massironi, Roland Wagner, and Katrin Stephan

13.1 Evolution of Asteroids and Dwarf Planets

The interplanetary space of our Solar System is not empty. During each starry night a stargazing observer may recognise a few shooting stars, called *meteors*. Those are tiny dust particles that enter Earth's atmosphere at velocities of several kilometres per second. During a single year periods of increased meteor activity may be recognised. At such times the Earth crosses streams in which the particle density is increased. In most cases the source of such streams are comets. As detailed in Sect. 13.2.4 they contain high fractions of volatile materials mixed with dust. As volatiles sublimate due to solar insulation, the dust is released and due to its moment of inertia it remains near the orbit of its source. Over time these dust trails are cleared out by solar radiation and the solar wind, if not topped up again by the source body. Occasionally very bright meteors (bolides) are observed. These are usually larger pieces of mostly rocky material that enter the atmosphere at similar velocities as average meteors. Pieces of bolides could reach the ground and may be picked up as meteorites that can be analysed in laboratories. Yet larger rocks from space cannot be slowed down by Earth's atmosphere efficiently enough and impact its surface with cosmic velocities. In such cases craters are formed and most of the meteoritic

N. Schmedemann (✉)
Freie Universität Berlin, Berlin, Germany
e-mail: nico.schmedemann@fu-berlin.de

M. Massironi
Universitá degli Studi di Padova, Padova, Italy
e-mail: matteo.massironi@unipd.it

R. Wagner • K. Stephan
German Aerospace Centre (DLR), Berlin, Germany
e-mail: roland.wagner@dlr.de; katrin.stephan@dlr.de

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_13

material is vaporised due to the high energy released during the impact event (see Chap. 6).

Most meteorites are stony or iron fragments of Main Belt asteroids released during collision events. They present evidence for the chemical and physical makeup of asteroids. A fraction of less than 0.1% of the meteorites derive from the Moon and about 0.4% from Mars. Due to their high volatile content, meteorites from cometary objects are not stable on Earth's surface and thus no meteorites are known from such objects.

Today, a growing number of small bodies are investigated up close by spacecraft (see Table 13.1), as the vast majority of minor bodies in the Solar System is spatially not yet resolved by telescopic observations. We know about their shapes, size and rotational state only by changes and absolute values of their *light curves* or in cases of near Earth flybys, resolving radar observations. If the light spectrum is analysed, the surface composition of the small bodies can be revealed as well. In some cases it is even possible to link specific meteorites found on Earth with their likely parent bodies. About 100–150 small bodies have been suggested as possible source of the

Table 13.1 List of small-body encounters by spacecraft; letter *P* indicates cometary bodies

Number	Name	Spacecraft	Min distance (km)	Date
5535	Annefrank	Stardust	3079	02 Nov 2002
132524	APL	New Horizons	101,867	13 Jun 2006
9969	Braille	Deep Space 1	26	29 Jul 1999
1	Ceres	Dawn	845	Since Dec 2015
433	Eros	NEAR Shoemaker	Landed	12 Feb 2001
951	Gaspra	Galileo	160	29 Oct 1991
26P/	Grigg–Skjellerup	Giotto	200	10 Jul 1992
1P/	Halley	Giotto	596	14 Mar 1986
		Sakigake		
		Suisei		
		Vega 1–2		
243	Ida	Galileo	2390	28 Jun 1993
25143	Itokawa	Hayabusa	Landed	20/25 Nov 2005
21	Lutetia	Rosetta	3170	10 Jul 2010
2685	Masursky	Cassini	1,600,000	23 Jan 2000
253	Mathilde	NEAR Shoemaker	1212	27 Jun 1997
134340	Pluto	New Horizons	12,500	14 Jul 2015
2867	Steins	Rosetta	800	05 Sep 2008
9P/	Tempel 1	Deep Impact		
		Stardust	181	15 Feb 2011
67P/	Tschurjumow–Gerassimenko	Rosetta, Philae	Landed	12 Nov 2014
4	Vesta	Dawn	465	13 Dec 2011
81P/	Wild 2	Stardust	240	02 Jan 2004

analysed meteorites but uncertainties are high, due to the large number of potential parent bodies.

Major concentrations of minor bodies can be found between the orbits of Mars and Jupiter, in the asteroid *Main Belt* and in an extended region beyond the orbit of Neptune, the so-called *Kuiper Belt*. Minor bodies such as asteroids and comets are considered to be the remains of the building blocks of planetary formation in our Solar System dating back to about 4.5 Ga ago. During that time they underwent a geologic development that is discussed in this chapter separately for the major groups of minor bodies.

While asteroids are known predominantly from the asteroid Main Belt, the cratering records of Jovian and Saturnian satellites suggests that asteroids are not confined to the mentioned regions but may be scattered by dynamical interactions with major bodies into the outer Solar System as well.

13.1.1 Formation

The Solar System formed from a cold interstellar cloud of gas (∼99%) and dust (∼1%) that collapsed under its own gravity as a result of destabilisation perhaps due to exposure to shockwaves of a nearby supernova explosion. The cloud likely fragmented into several regions of increased density at which centres new stars were born, surrounded by protoplanetary discs. One of those stars was our Sun. Such discs have been observed for example within the famous Orion nebula (M42, Fig. 13.1a).

The dust particles of the protoplanetary discs *agglomerate* due to electrostatic forces into larger particles within a relatively short amount of time. When the agglomerates become sufficiently massive gravitational attraction is going to be the

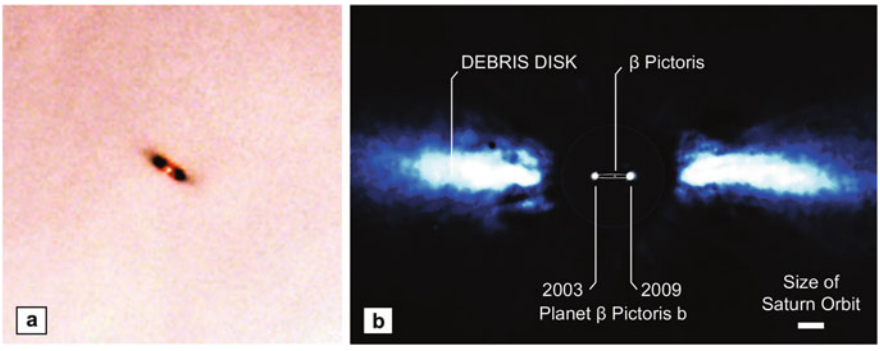

Fig. 13.1 (**a**) Protoplanetary disc around a young star in the Orion nebula, M42 (source: NASA/ESA and L. Ricci, ESO); (**b**) Young star system Beta Pictoris (age: 10–30 Ma) with debris disc and a planet with ∼8 times the mass of Jupiter at ∼9 AU distance from its host star Beta Pictoris (source: modified from ESO1024—Science Release)

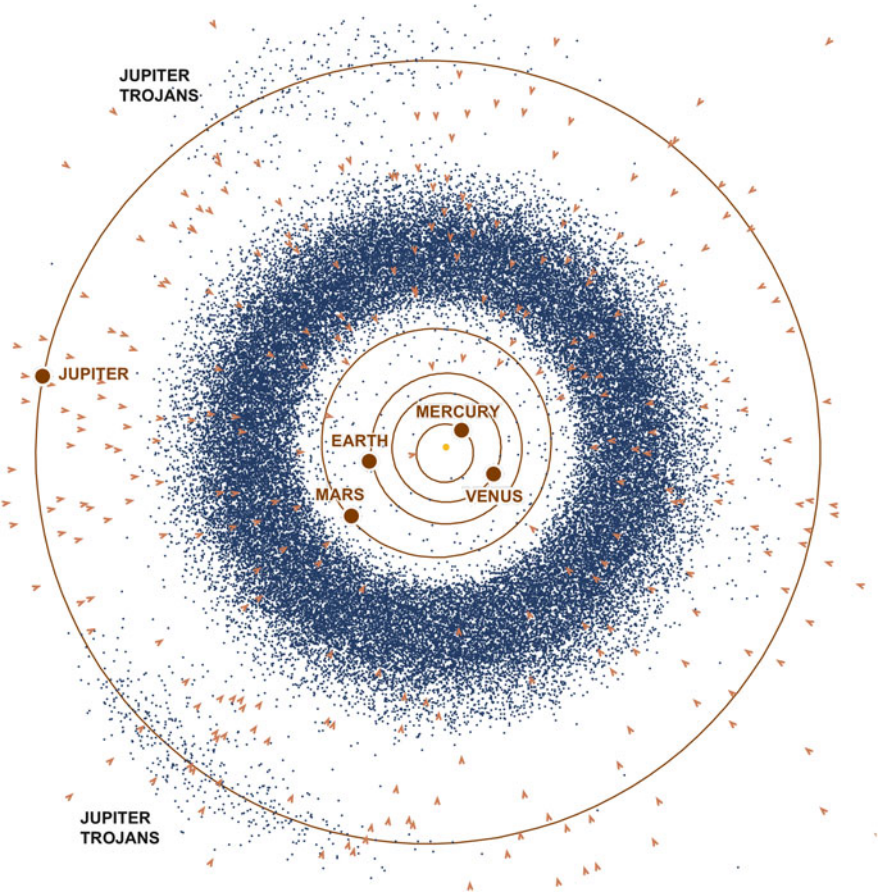

Fig. 13.2 Distribution and scale of small bodies in the inner Solar System (see also 1.5 on page 10, after P. Chodas (NASA/JPL))

dominant force that result in a runaway *accretion* process. Within only a few tens of millions of years major planets are formed and the particle areal density of the disc is significantly reduced (Fig. 13.1b). In our own Solar System gravitational interaction of the major planets prevented the formation of another large planet between the orbits of Mars and Jupiter. That can be inferred from the large number (80% of all falls) of chondritic meteorites that are derived from Main Belt asteroids (Fig. 13.2).

Chondritic material was never part of a planetary body large enough to produce heat in excess of the solidus temperature of the silicate material the chondrites contain. Chondrites contain *Calcium-Aluminium inclusions (CAI)* which is believed to be the oldest (4.567 Ga) material of the Solar System. Chondrules that are also found in chondrites are small spheres that formed from dust in the early solar disc during a very short but intense heating event about three million years after

CAIs (see Chap. 6). This event could possibly have been a supernova explosion that enriched the solar nebula with radioactive material. Products of radioactive decay of short-lived isotopes typical for supernova explosions (e.g. ^{60}Fe, ^{26}Al) were found in many different classes of meteoritic material. Recent work suggests that instead of a supernova explosion strong stellar winds of massive stars (e.g. Wolf-Rayet stars) that formed in the Suns stellar nursery may be responsible for high influx of specifically ^{26}Al to the young solar protoplanetary disc.

Excluding meteorites from the Moon or Mars, there is still about 20% of meteoritic material that underwent a process of differentiation within larger bodies. Such meteorites are called *achondrites*. For instance about 6% of all found meteorites are likely derived from the basaltic crust of the large (∼520 km diameter) asteroid (4) Vesta. Vesta probably formed within the first five million years after CAIs (Chap. 6) and due to its size incorporated significant amounts of short-lived radioactive isotopes. Radioactive heating of Vesta supported its differentiation into a metallic core (∼220 km diameter), olivine-rich mantle and a basaltic crust. Vesta is the only large differentiated asteroid in the Main Belt that is still intact. However, iron meteorites (∼4% of all falls) are pieces of other differentiated small bodies that were catastrophically disrupted, leaving their metallic cores exposed to continued bombardment by other asteroids. One example, (16) Psyche, is an 180 km diameter iron core of a former differentiated asteroid, very similar in size to Vesta. In comparison to iron meteorites one might expect to find many more meteorites that derive from the mantle of differentiated asteroids. But instead only a small fraction of meteorites, rich in olivine, can be linked to that region, which is an unsolved problem yet. For instance pallasites derive from the core/mantle boundary and are rich in olivine and metallic iron (see Chap. 6).

13.1.2 Composition

In general the bodies of the Solar System formed from the same cloud of gas and dust like the Sun but have been altered by various processes e.g. differentiation and subsequent redistribution within the Solar System. The elemental composition of almost unaltered carbonaceous chondrites thus, represents the composition of the dust phase of the interstellar cloud from which the Solar System formed and which is similar to the solar composition neglecting highly volatile elements (H, N, C and the noble gases). The heavier elements formed mineral grains which in turn are the construction blocks of the rocky material that constitutes asteroids.

Subsequent differentiation and redistribution might change the elemental composition of asteroids depending on their environment and geologic evolution. Thus, mineralogy changes according to the nature and extent of aforementioned processes. From minerals on Earth as well as those found inside meteorites there is a good

understanding about the spectral characteristics of the different types of minerals. This knowledge is used to identify the mineralogic composition of asteroids by accurate analysis of the spectral characteristics of the light received from the asteroids.

Asteroids can be taxonomically classified by their specific spectral characteristics. The *Tholen* classification scheme includes 14 different types that are defined by eight spectral bands ranging from 0.31 to 1.06 μm. The more precise *Small Main-Belt Asteroid Spectroscopic Survey* (SMASS) classification is oriented at the Tholen scheme but includes 24 different classes due to higher spectral resolution in a slightly narrower range from 0.44 to 0.92 μm. Most asteroids belong to the C-group, S-type and X-group.

- The C-group constitutes of relatively dark carbonaceous objects with high fractions of carbonaceous materials, organic material even amino acids, phyllosilicates, Olivine, Pyroxene, metals and volatiles such as water ice. The C-group contains the Tholen types B, C, F and G. Examples for this group are (2) Pallas (B-type), (10) Hygiea (C-type), (13) Egeria (G-type) and (228) Agathe (F-type). C-group asteroids may be the parent bodies of the most primitive carbonaceous chondrites.
- The S-type asteroids are of silicate or stony nature and contain less volatiles than C-types. Examples for S-type asteroids are (243) Ida and (951) Gaspra, which both have been visited by the Galileo spacecraft (see Table 13.1). S-type asteroids may be the source of ordinary chondrites.
- X-group asteroids show similar spectra within their group, but may be of different composition. X-group types E and especially M contain asteroids with high abundances of metal-rich minerals such as Enstatite. Examples for E and M type asteroids are (44) Nysa (E-type) and (16) Psyche (M-type). The E-type asteroid (3103) Eger may be the parent body of Aubrite meteorites, while M-types may be the parent bodies of iron meteorites that clearly underwent differentiation (Chap. 6). Such asteroids are therefore highly evolved. The X-group type P on the other hand appears to contain primitive organic rich silicatic asteroids with indications for carbon as well as water ice. An example for a P-type is (46) Hestia. The density of Hestia is estimated with more than 5000 kg/m^3 which would still imply a high fraction of metals.

Within the asteroid Main Belt the different classes have a tendency to populate different regions (Fig. 13.3) with more S-type asteroids at smaller semi-major axis and more C-group asteroids at larger semi-major axis. X-group bodies scatter throughout the Main Belt. The increasing amount of the low albedo carbonaceous C-group asteroids towards the outer Main Belt may also be responsible for the respectively decreasing average albedo of the Main Belt bodies (Fig. 13.3).

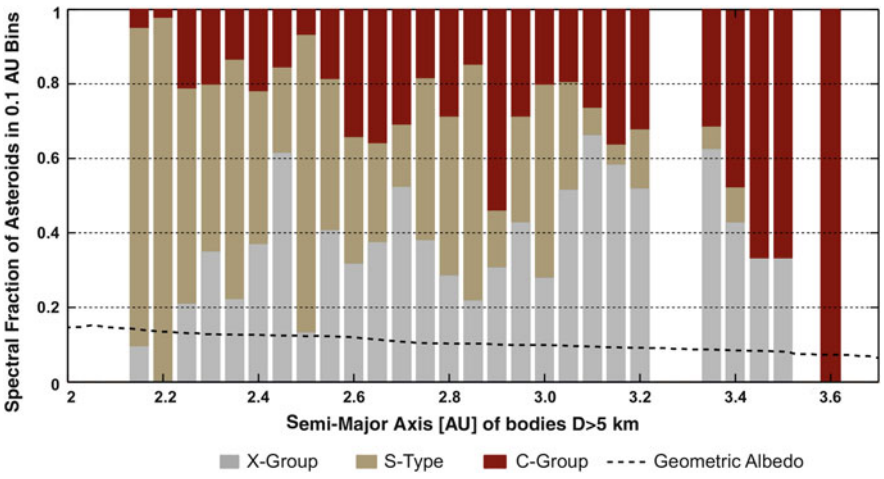

Fig. 13.3 Spatial distribution of the three major spectral groups of Main-Belt asteroids and respective average albedo of all Main Belt asteroids with known albedo

13.1.3 Dynamics

The asteroid Main Belt is—and it has probably been since the very early Solar System—heavily disturbed by gravitational perturbations of the major planets, first and foremost by Jupiter. These perturbations prevented the formation of a major planet in this region. Today the mass of the whole asteroid Main Belt is only ∼1% that of the Moon.

According to lunar cratering rates over the last ∼4 Ga and estimates on the surface density of the protoplanetary disk of the Solar System, however, the initial mass in the region of the asteroid Main Belt was several hundreds, if not a thousand times higher than it is today. Most of the mass was probably depleted already when the major planets were still forming, since Mars is part of the mass deficiency in the Main Belt region. During this time *Mean Motion Resonances* (*MMR,* Fig. 13.4) of the major planets were likely sweeping through the Main Belt because the conservation of angular momentum of ejected small bodies forced the major bodies to shift their orbits accordingly. In this period of time it may have happened that major planets ran into a mutual resonance catastrophe that caused a relatively short but intense release of small bodies that predominantly originated at the inner edge of the current Main Belt and bombarded the inner Solar System. This event is also known as *Late Heavy Bombardment* or *terminal lunar cataclysm* (see Chaps. 7 and 11).

Whether this event actually happened or not is debated since radiometric ages of lunar rock samples from the Apollo and Luna missions became available. However, there is general agreement that the currently observed projectile flux into the inner Solar System is constant within a factor of about 2 since ∼3 Ga. Based

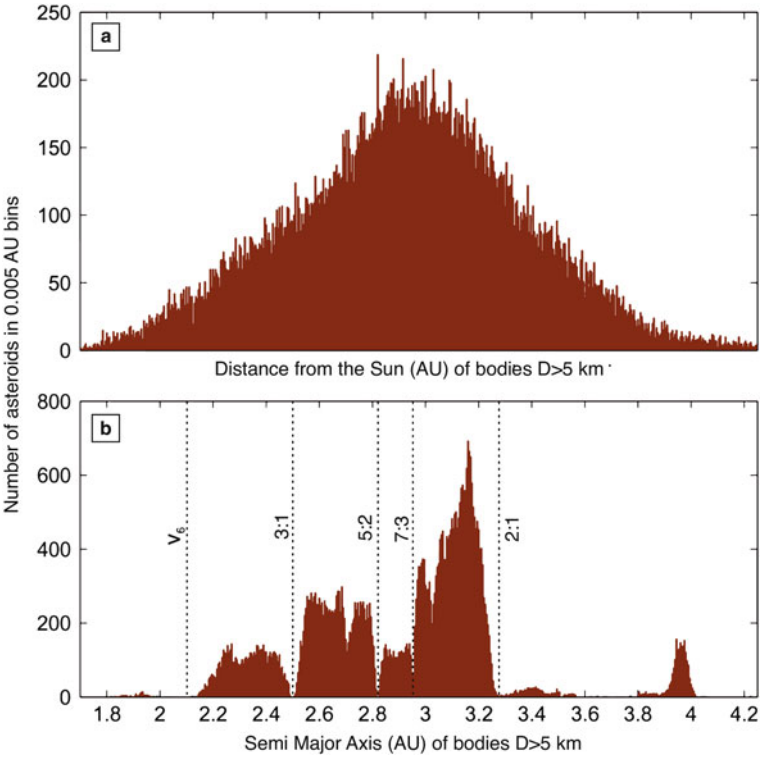

Fig. 13.4 (**a**) Distribution of distances from the Sun for asteroids >5 km at a specific point in time (November 2015); (**b**) semi-major axis of the same asteroids from the *upper panel*; binned number of asteroids shows multiple dips indicative for MMRs. ν_6 is a secular resonance of the Perihel positions of the asteroids and the one of Saturn. All other shown resonances stem from even number ratios of the orbital periods of the asteroids and Jupiter. Many weaker resonances exist due to interactions with the other major planets

on comparisons of the body size-frequency distribution (SFD) of near Earth objects (NEOs), the crater SFD of inner Solar System bodies such as the Moon, Mercury or Mars and the Main Belt body SFD, it appears likely that predominately Main Belt asteroids bombarded the inner Solar System bodies (see Chap. 7).

In the current dynamical setup of the Solar System there are about 700 *Near-Earth Objects (NEO)* $\geq$1 km diameter, which cross the orbits of the major planets and will impact them eventually. This transient projectile population is refilled from the Asteroid Main Belt by chaotic dynamical processes that involve multiple MMR of the major planets within the Belt as well as non-gravitational forces such as the YORP- and Yarkovsky-effects. The YORP effect (Yarkovsky–O'Keefe–Radzievskii–Paddack) is caused by interaction of surface material properties, topography, tilt of the rotation axis, respective illumination condition and the state of rotation of a small body. The Yarkovsky effect leads to a body migration in semi-major axis depending on previously mentioned material and rotational properties,

in such a way that the body slowly drifts inward if it rotates retrograde and outward if it rotates prograde. As the ratio of the surface area and volume or mass of the body increases towards smaller bodies, they are more affected by such radiation forces as larger bodies. Therefore, in particular small bodies are drifting around in the Main Belt, impact on other asteroids or eventually drift into one of the numerous resonance zones from which they are exited onto highly eccentric planet-crossing orbits.

Once a small body crosses the orbit of a major planet, there is a high probability of being scattered onto a significantly different orbit and thus of changing its semi-major axis. At any specific point in time the MMR are not visible from the actual positions of the asteroids, because many of them follow orbits with eccentricities around 0.1 that are much larger than those of the major planets. Thus, asteroids spend most of the time relatively far from their semi-major axis. The distribution of asteroids shows a peak around 3 AU but no sign of MMRs. But if semi-major axis of the same bodies are plotted the binned number of Asteroids does show multiple dips that result from orbital destabilization due to MMRs. Dynamical simulations can be used in order to follow the orbital evolution of specific asteroids when they get excited. Asteroids can leave the Main Belt due to excitation and scattering by gravitational interaction with the major planets.

The orbital evolution of the asteroids starts inside the Main Belt with relatively low eccentricity ($\sim$0.1). Once the asteroids drift into a MMR they get dynamically excited by increasing their eccentricities at constant semi-major axis. At sufficiently high eccentricities they become planet-crossing asteroids with Mars and/or Jupiter. Because Jupiter has more mass than the small inner Solar System bodies, it influences the orbits of the small bodies more efficiently than other planets. If an asteroid is gravitationally scattered for example by Jupiter, this happens relatively close to Jupiter itself and thus to its orbit. The point where the scattering happens is part of the new orbit of the asteroid which will therefore visit Jupiter on a regular basis. At sufficiently high eccentricities such a member of the Jupiter family is still able to cross the orbit of other major planets where a gravitational scatter could happen again. Thus, some of the asteroids are able to become members of the dynamical family of other planets as well.

13.1.4 Geological Evolution

The geological evolution of most asteroids is dominated by collision events with other asteroids. Shortly after formation of the Solar System when short-lived radioactive materials were still abundant in the rocky material, larger asteroids experienced limited endogenic geologic evolution. For instance the parent bodies of iron meteorites differentiated similar to the major planets into a metallic core, an iron depleted mantle and basaltic crust. Only one of those bodies survived the heavy bombardment in the Main Belt until today. (4) Vesta was selected as target for the Dawn mission in order to understand the early stages of planetary formation

and it appeared to be an intact protoplanet. Since small bodies are losing their internal heat fueled by radioactive decay much more efficiently than the major planets, differentiation of asteroids probably stopped within 100 Ma after CAIs. Furthermore, it is likely that the growth of planetesimals was faster in the inner parts of the solar disk due to its higher surface density. Thus, basaltic asteroids such as Vesta probably formed only within ~2 AU from the Sun and were scattered into the Main Belt later. From fully differentiated asteroids like Vesta to asteroids that never experienced any metamorphism there are samples of intermediate stages of partial melting and metamorphism in our meteorite collections. Asteroids rich in volatiles may have differentiated into an anhydrous core a hydrated mantle and an icy crust, while the rock fraction did not differentiate.

Such models are discussed for the dwarf planet (1) Ceres that is also a member of the Main Belt but is discussed later. Such bodies show spectral evidence for aqueous alteration of rocky materials. Aside from endogenic geologic evolution, asteroids also experienced intense exogenic geologic evolution due to collisions with other asteroids. Many of the originally formed bodies were destroyed by catastrophic collisions very early in Solar System history. The frequency of mutual collisions among Main Belt asteroids is proportional to the number of potential projectiles. Thus, over the age of the Solar System collision frequencies dropped by 2 or 3 orders of magnitude. A result of the mutual collisions is the asteroid body size-frequency distribution (SFD). Depicted in relative representation it shows a characteristic W-like shape. Going from the inner to the outer Main Belt the amplitude of the wave between ~5 km (maximum) and ~25 km (minimum) is shrinking but the positions of minimum and maximum are not changing (Fig. 13.5a). Current collision probabilities of asteroids are dropping going from the inner Belt outwards. Thus, it appears reasonable to suspect that the inner Main Belt is collisionally more evolved than the outer Main Belt (Fig. 13.5b).

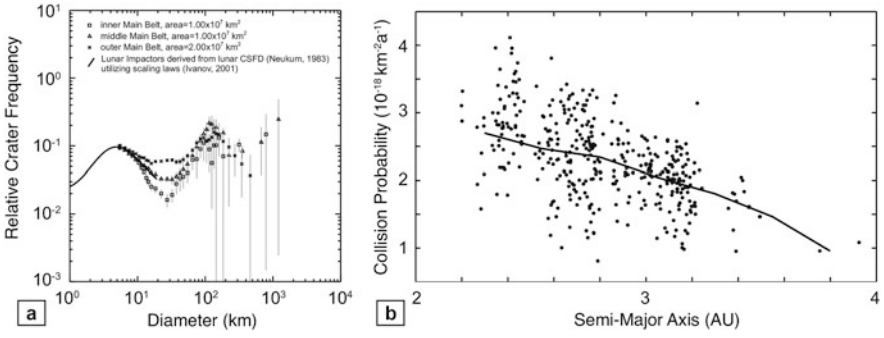

Fig. 13.5 (**a**) Relative size-frequency distribution of bodies of the inner, middle and outer Main Belt. Frequencies of bodies around 25 km diameter are lower in the inner Belt than in the outer Belt, if compared to frequencies at ~5 km; (**b**) *black line*: floating average of intrinsic collision probability for about 2200 Main Belt asteroids (*dots*) larger than 20 km. Collision probabilities among the Main Belt asteroids appear to decrease with increasing semi major axis

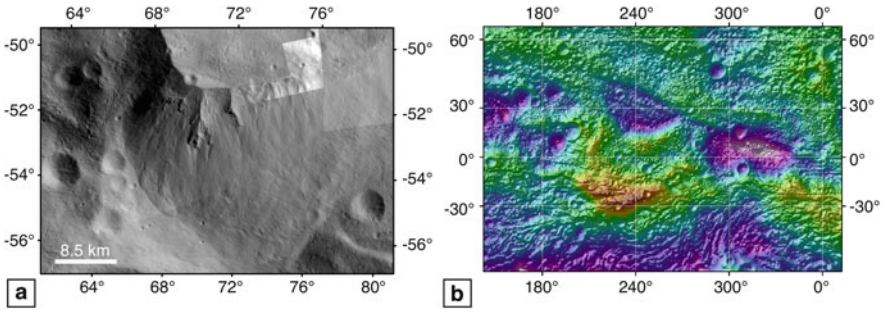

Fig. 13.6 (**a**) Southern tip of Matronalia Rupes on Vesta. It is a steep cliff within the rim of the Rheasilvia basin. Low crater frequencies in comparison to adjacent areas immediately indicate relatively recent mass wasting activity on the steepest parts of the cliff. (**b**) Hemispheric view of Vesta as color coded topography draped over a shaded relief model in Mercator projection. The *color* indicates heights between −20 km (*light purple*) to +15 km (*red-orange*) relative to an ellipsoidal reference body. Near 240°E/30°N is the centre of the ∼180 km Postumia crater which southern rim is well defined but not its northern rim. The crater is crosscut by a topographic step in NW–SE direction. Several trough-like features run parallel next to the step on its southern side. The troughs are named Saturnalia Fossae and are a tectonic expression of the formation of the Veneneia basin on Vesta. The younger and more massive Rheasilvia impact likely reactivated the existing fault system of Saturnalia Fossae and lifted the northern part of Postumia, thus muting the topographic expression of the northern crater rim. Wavy features near the image bottom indicate partly Coriolis deflected mass wasting into the Rheasilvia basin

Another result of the intense collisional history as well as the relatively low surface gravity of the asteroids is their highly irregular shape that is causing high topography on many Main Belt asteroids. For instance on Ida parts of the true surface are almost a factor of two above and below the surface of the reference body. This huge deviation has to be considered if any measurement is taken that is related to the reference body. In a more geological sense high standing topography is responsible for mass wasting processes and can cause extension cracks on the surface in the vicinity of large cliffs. Figure 13.6 gives an example for mass wasting on 4 Vesta.

However the majority of cracks visible on the surface of asteroids are probably related predominantly to significant impacts, such as the equatorial troughs on Vesta are a result of the huge Rheasilvia impact near the vestan South Pole or the inclined Saturnalia Fossae trough system that is tectonically related to the formation of the older and smaller Veneneia basin. On the other hand there are examples where prominent trough systems defy any explanation thus far. Probably the most prominent example for yet ambiguously explained cracks are the grooves on the Marsian satellite Phobos, which could be a captured asteroid.

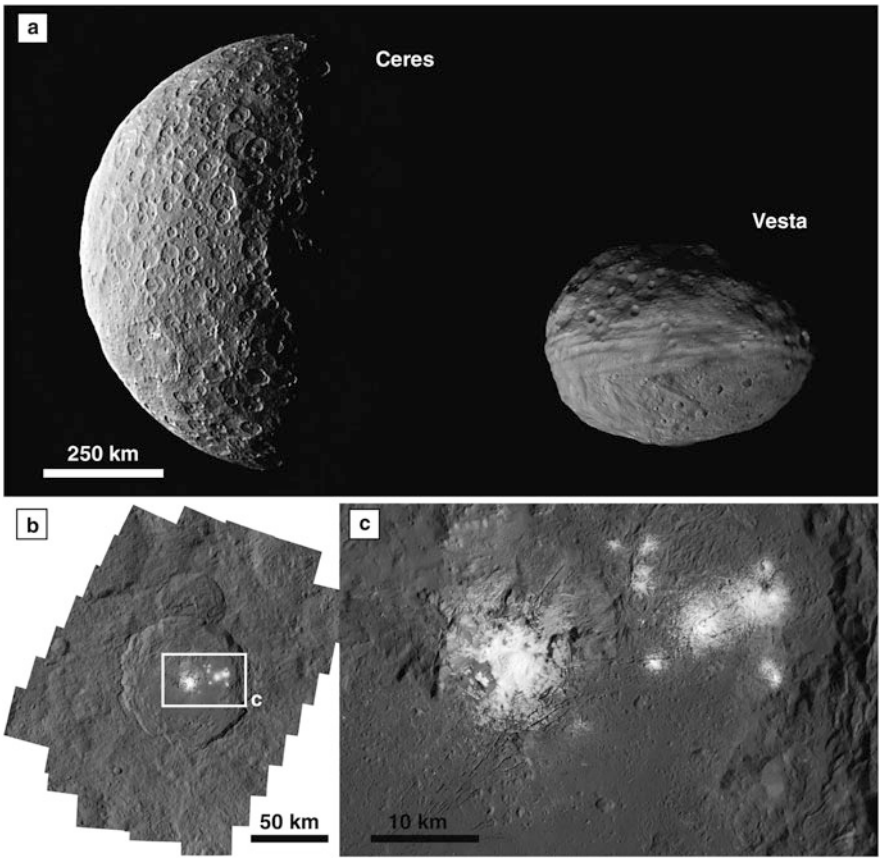

Fig. 13.7 A comparison of Ceres and Vesta; (**a**) Sizes if Ceres and Vesta compared; (**b**) Occator crater on Ceres with bright spots of recently deposited material; (**c**) close up of bright spots inside Occator. Source: (**a**, **b**, **c**) NASA/JPL-Caltech, UCLA, MPS, DLR, IDA

The relative size of Ceres and Vesta is displayed in Fig. 13.7. On Vesta patches of dark material were discovered that appear to originate from the projectile that formed the older (Veneneia) of the two South Pole basins (Fig. 13.8). As also shown in Fig. 13.8 the area with lower albedo in the vicinity of the Veneneia basin is also characterized by relatively high abundances of hydrogen. Thus, the surficial composition of asteroids is a result of contamination with exogenic material, impact gardening as well as space weathering that is changing color and albedo properties of the regolith for instance due to micro meteorite bombardment and destruction of the crystallographic structure of minerals from cosmic radiation. The resulting global geology of Vesta is thus relatively complex, as displayed in Fig. 13.9.

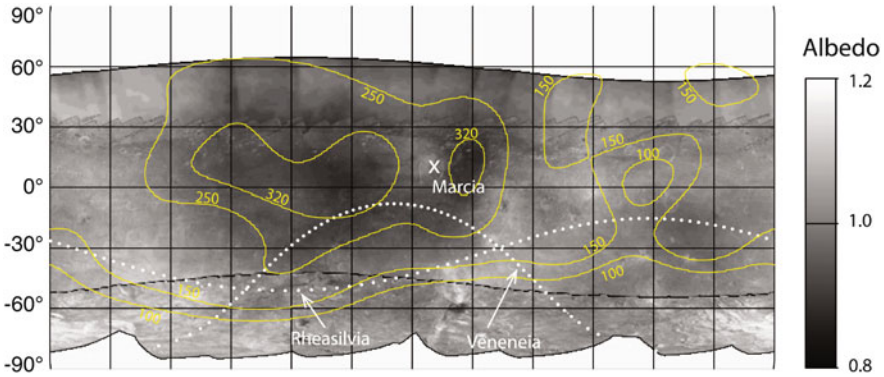

Fig. 13.8 Albedo map of Vesta with hydrogen abundance (*yellow* contours). *Dotted curves* represent the rims of the Rheasilvia and Veneneia basins. Highest hydrogen abundances are measured in a low albedo area north of the Veneneia basin (central meridian at 180°, JPL Photojournal - PIA 16181)

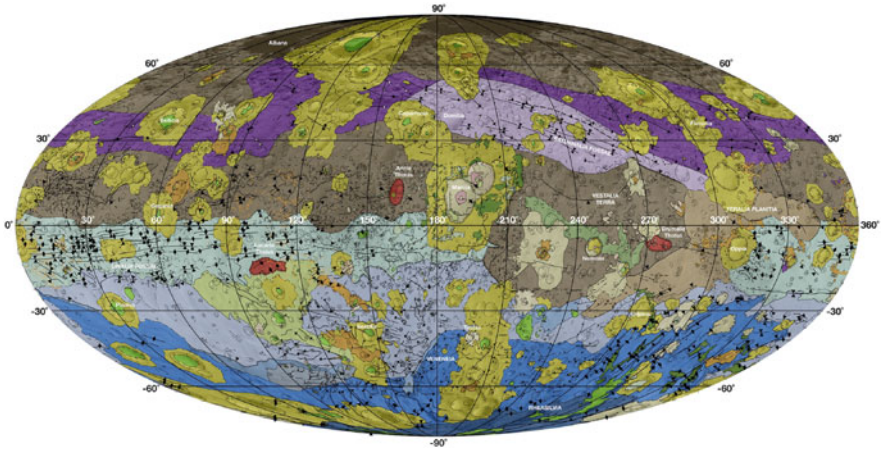

Fig. 13.9 Geological map of Vesta. Source: USGS, NASA/JPL-Caltech/ASU

Ceres (Fig. 13.7a) is much larger and displays a wider range of geological features, in addition to retaining a rich impact record on its heavily cratered surface (Chap. 7): its crust with a topographic range of several kilometers suffers deformation from both impact and tectonic processes. One of the notable features on Ceres are central bulges in few impact craters, cracked and associated with bright deposits, such as Occator (Fig. 13.7b, c). Those might be related to subsurface volatile reservoirs and even possibly a subsurface ocean with some communication with the surface. The overall crust might contain a large amount of volatiles itself, but not likely pure ice, based on the number and morphology of retained craters, different from those of icy satellites (Chap. 12). The global geology of Ceres is being mapped.

Dwarf planet Pluto has been a target for planetary astronomy rather than geological sciences for the last several decades, yet since the recent flyby of NASA New Horizons its (only partially expected) rich and complex surface geology has been revealed: pre-flyby hyphotheses on Pluto's geology were mainly based on analogies with icy satellites in the Jupiter and Saturn systems (see Chap. 12), which turned out to be rather good, although actual data surpassed expectations, as often is the case.

The surface of Pluto is composed of ices such as N_2, CH_4 and CO_2 at a temperature of about 40 K. Those volatiles contribute to maintaining the atmosphere, currently not massive enough to support liquid phases of those ices, but possibly at some point in the past with high enough pressure to support palaeolakes, based on shoreline-like features. Likewise, dissection of terrains with drainage-like morphologies exists on Pluto's highlands. The dwarf planet retains its impact record on large portions of its surface, dating back up to about 4 Gyr ago, while large areas have been resurfaced extremely recently, possibly in geological present times (Fig. 13.10), with crater-based ages of less than 10 Myr.

The global geology of Pluto is in fact dominated by the informally named Sputnik Planum, a roughly heart-shaped large depression (Fig. 13.10a) surrounded by rugged, eroded terrains with topographic rise of several km above the plain. The interface between those terrains is characterised by the presence of chaotic terrains, a type of surface feature shared, across very different geological settings, by diverse bodies such as Europa or Mars (see Chaps. 11 and 12). In this case it is linked to remobilised, disrupted ice blocks. Around Sputnik Planum glacial flow features exist while its interior displays evidences of convection, with cell-resembling surface patterns (Fig. 13.10b). The results of cryovolcanic processes, as expected, are largely visible on Pluto and also large-scale cryovolcanic edifices are present (Fig. 13.10c). Pluto's moon Charon has a remarkable geological diversity and a degree of activity, including global tectonic deformation, particularly across a low latitude bands, which also acts as a dichotomy separating hemispheres with heterogeneous geology (see Chap. 12).

13.2 Evolution of Comets

Comets are thought to have fertilized the Inner Solar System delivering the organic material pivotal for the origin of life. They, together with chondritic meteorites, are the only bodies that can provide the ground truth for understanding the physical and chemical processes in the Solar Nebula that led to our planetary system. Being icy, planetesimals formed in colder region of the Solar Nebula (beyond the snow-line) and almost unaffected by gravitational compression because of their small size and low density, they may have retained the volatile compounds as well as the primordial refractory materials out of which the Solar System was built. However, although chemical composition of comets can readily match the primordial accreting material within the proto-planetary disk, their nuclei morphology can either reflect the early

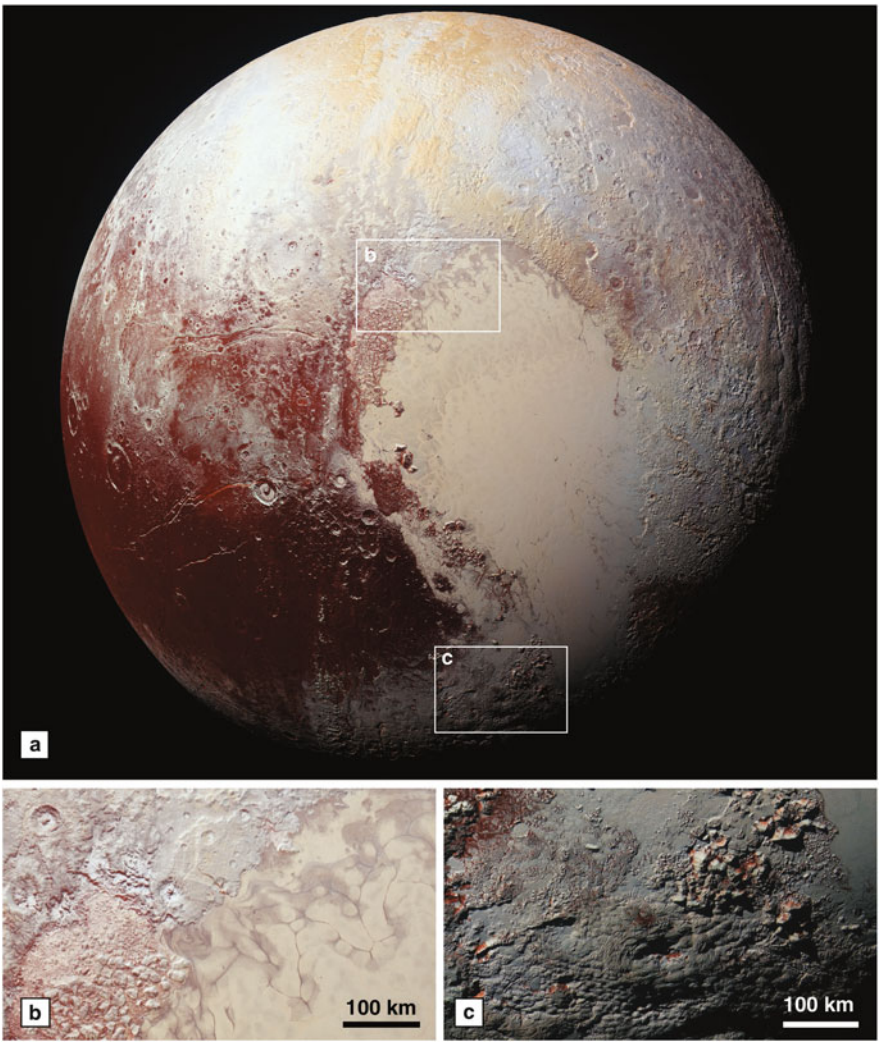

Fig. 13.10 Pluto as imaged during the NASA New Horizons flyby, in enhanced colors, blue, red, near infrared. The main physiographic and geologic provinces are visible, as well the global albedo variations; (**a**) almost complete disk in color, displaying the diverse surface Geology, ranging from heavily cratered terrains to virtually craterless ones; (**b**) interface between eroded highlands and bright lowlands (Sputnik Planum), marked knobs in chaotic terrains; (**c**) large mound, informally named Wright Mons: It is likely to be a cryovolcanic edifice, hundreds of km across, located at the centre of the image and surrounded by rugged terrain. Source: NASA/Johns Hopkins University Applied Physics Laboratory/Southwest Research Institute

accretionary processes or be affected, at various degree, by collisional events, space weathering, intense sublimation and erosive processes. Hence comets might have preserved the imprints of the early material accretion from the Solar Nebula and by all means can be considered fossils of the nascent Solar System (Fig. 13.1a, b).

13.2.1 Orbits and Reservoirs

As seen from Earth comets can be very beautiful with spectacular comas and tails or mediocre with weak comas and small tails, invisible to the naked eye. The former are generally the long-period comets which are characterized by periods often exceeding 200 years up to about 10 Ma, orbits inclination varying widely over the full range and aphelia far beyond the external planets. Due to their orbital characteristics, JAN HENDRIK OORT (1900–1992) proposed that the reservoir of long-period comets is a spherical cloud of up to 10^{13} small objects extending from about 10^3 to about 10^5 AU (see Chap. 1). Passing stars, which have relative motion one to each others, and giant molecular clouds in the interstellar medium can perturb the Oort Cloud objects throwing some of them out of the Solar System or reducing the perihelion distance of others to values that enable coma and tails to form. Galactic tides induced by differential gravitational forces of stars and interstellar matter can cause similar effects.

Comets with less remarkable tails and comas commonly belong to the short-period group and are characterized by a periodicity shorter than 200 years, semi-major orbital axes less than 34 AU and low orbital inclinations (mostly less than 35°). Most of them called *Jupiter family comets* have periods lower than 20 years, perihelion distances of one to a few AU and aphelions in the Jupiter region. Although these comets are affected by frequent orbit alteration due to close encounters with planets, their lower orbital period make them the preferred objects for cometary space missions (see Table 13.1). Due to their low orbital inclination and shorter semimajor axis the proposed reservoir for short period comets is a belt extending beyond the orbit of Neptune to about 45 AU. The belt is named after KENNETH E. EDGEWORTH (1880–1972) and GERARD P. KUIPER (1905–1973) who first proposed it. In-between the *Edgeworth-Kuiper Belt* and the *Oort Cloud* is the transitional *Scattered Disk*, which extends from about 45 to 10^3 AU. It is estimated that 10^{10} objects should populate the short-period comet reservoir. The perturbations needed for converting these objects into short-period comets are guaranteed by the gravitational interactions with the giant planets (Neptune in particular) and favored by collisions among the Kuiper belt objects themselves.

13.2.2 Origin

Comets formed in the nascent Solar System beyond the snow line of the proto-planetary Solar Nebula. The snow line is defined as the site beyond which the water condenses and it is placed at about 5 AU from the proto-Sun. Hence comets originated in the cold outer Solar System where giant embryos formed leaving behind a large number of icy-rocky planetesimals. The volatile contents of such planetesimals was related to their place of formation being the Jupiter-Saturn region warmer (lower contents of volatiles) than the Uranus-Neptune one (higher contents of volatiles). The planetesimals which interacted with Jupiter or Saturn were captured or thrown out of the Solar System, whereas the less energetic interactions with Uranus and Neptune fed the Oort Cloud with volatile rich rocky-ice objects. Beyond Neptune the increasing heliocentric distances of the planetesimals prevented the formation of evolved bodies and a population of small icy-rich objects remained where they formed constituting the Edgeworth-Kuiper belt and Scattered Disk reservoirs. The planetesimals stored in the cold outer Solar System reservoirs remained potentially unaffected by major modifications for 4.5 Ga until to their deflection into the present orbits with perihelion distances within about 6 AU and consequently more or less relevant activation.

One of the major debates in Solar System science is whether cometary nuclei are relatively unprocessed aggregates directly formed from the proto-planetary Solar Nebula (primordial rubble-piles) or welded collisional debris from larger parent bodies (collisional rubble-piles). In the first case comets would be really old objects which may give us insights into the agglomeration processes that transformed the granular proto-planetary disk into a planetary system; in the second case, they could have undergone significant structural and superficial changes due to collisions, even in the Edgeworth-Kuiper Belt or the Scattered Disk, and teach us more about the physics of collisional disruptions and gravitational re-junction than about the chemical, mineralogical and physical properties of the Solar Nebula. The ESA/Rosetta spacecraft, orbiting around comet 67P/Churyumov-Gerasimenko (67P/CG) between August 2014 and September 2016, has provided a wealth of data about the physics, composition and morphology of a cometary nucleus that seems to favor a primordial origin rather than a collisional one. These comprehend its extreme low density and high porosity (similarly to all the other cometary nuclei), the relevant content of supervolatiles (CO and CO_2) and the evidence of a thick and extensive stratification.

13.2.3 Overall Anatomy and Fate

When comets approach the Sun at heliocentric distances classically lower than 2.5–3 AU, the volatile compounds (H_2O, CO and CO_2) of their nucleus start degassing giving to comets their renown external anatomy composed by a head and, generally, two long tails (Fig. 13.11a, b).

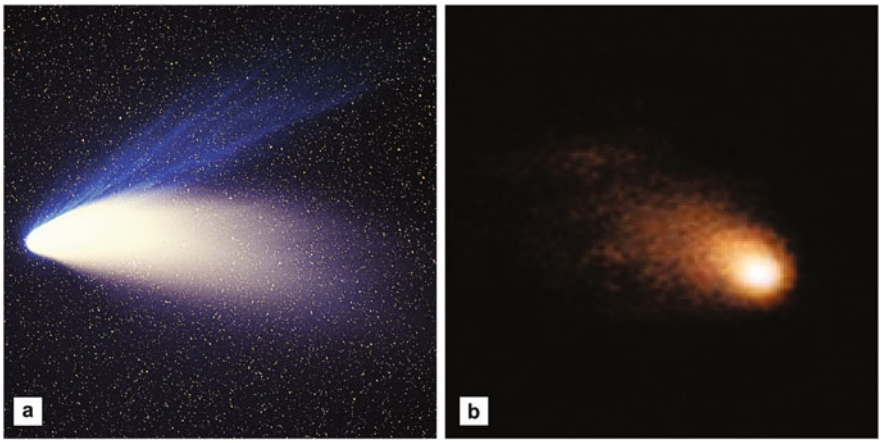

Fig. 13.11 Comets Hale Bopp (C/1995 O1) and 67P/Churyumov-Gerasimenko (67P/CG); (**a**) long-period comet Hale-Bopp, note the two long tails: the *bluish* ion tail which point away and radially from the Sun and the dust tail, *yellowish* and arched; (**b**) Short-period comet 67P/Churyumov-Gerasimenko (67P/CG) view from the European Southern Observatory's Very Large Telescope in Chile on 11 August 2014. Source: (**a**) E. Kolmhofer, H. Raab; Johannes-Kepler-Observatory, Linz, Austria. (**b**) European Southern Observatory's Very Large Telescope (ESO/VLT)

The coma is the visible part of the head composed of a bright cloud of gas and dust roughly of 10^4–10^5 km of diameter. It is surrounded by a wide hydrogen cloud, which is generated by dissociation of H-bearing molecules under the effect of UV solar radiation. Hidden within the coma is a few kilometer to tens of kilometers large nucleus which is of major interest for planetary geologists since the emission centres of sublimating gas and ejected particles (jets) are located at its surface. Nuclei are not visible with any Earth based observation and can be explored only through close encounters. After the Giotto and Vega missions, which visited Halley comet in 1986 and imaged for the first time a comet nucleus, other three missions were dedicated to cometary nuclei the most recent being ESA Rosetta/Philae (see Table 13.1). From these missions we directly confirmed that cometary nuclei are extremely low density bodies (500 kg/m^3 on average) composed by a mixture of ice and dust with a porosity around 70–80%.

The ion tail is generated by the solar radiation through photodissociation and photoionization of the sublimating gas molecules emitted by the comet. Together with the more common OH^+ and H_2O^+ ions, the presence of CO^+, N_2^+ and CO_2^+ make the ion tail glowing in the dark with bluish colors since these ions are characterized by an electromagnetic emission in this wavelength range. Cometary ions are swept away from the solar wind with a speed much larger than the orbital one of the comet, hence the ion tail is straight, point away and radially from the Sun and might extend up to 10^8 km from the comet head.

The dust tails are made up of both dust and gasses with a mass ratio between 0.1 and 10. The dominant small dust particles (down to 0.1 μm in size) reflect the sunlight in the visible wavelength range and are pushed outward from the solar radiation with respect to the nucleus orbit. This leads to a lag between the slower dust particles at higher heliocentric distance and the faster nucleus at a lower distance form the Sun. The result is a characteristic curved yellowish tail up to 10^7 km long.

The long-period comets, such as Hale Bopp and Hyakutake, are characterized by spectacular tails since they rarely pass close to the sun and their nuclei are not yet covered by degassing leftovers of refractory materials from previous perihelion passages (Fig. 13.11a). On the contrary, short-period comets have weaker tails since the activity of the nucleus is often hampered by a mantle of refractory carbonaceous and silicate dust particles covering most of its surface (Fig. 13.11b). Ultimately this mantle can become so continuous and thick that the nucleus is not any more able of generating coma and tails. Comets can have much more stunning ends than that, colliding with other planetary bodies and, more often, being fragmented and inexorably attracted by the gravitational pull of the Sun or Jupiter. The most outstanding example of this case is the fragmentation and collision with Jupiter of the D/Shoemaker–Levy 9 comet in 1994.

13.2.4 Composition

The composition of cometary nuclei can be indirectly derived from ground-based remote-sensing investigation of spectral lines and in-situ observations through spacecraft close encounters. The Giotto and Vega missions to comet 1P/Halley measured comas composition using mass spectrometers in 1996. The Stardust mission collected grains from comet 81P/Wild in 2004 and returned them back to the Earth for analyses. The Deep Impact spacecraft released an impactor, which hit the 9P/Tempel 1 nucleus triggering material emission remotely analyzed by the fly-by spacecraft itself, the Rosetta spacecraft, Hubble Space Telescope and ground based telescopes. Finally a great part of the Rosetta orbiter payload was dedicated to the compositional analysis of 67P/CG coma and nucleus (e.g. mass spectrometer, imaging spectrometer, grain collectors). Even the landing probe Philae was equipped by two mass spectrometers that were able to collect and analyze dust and gasses released at its first touch down on comet 67P/CG.

13.2.4.1 Volatiles

The most abundant molecules of comas are H_2O, CO and CO_2, that typically cover 95% of the total gas density. CO and CO_2 are generally referred as super-volatiles being released at temperatures of 25 and 80 K respectively, much less than that of water (180 K). Generally H_2O molecules largely dominate over the others, but on 67P/CG diurnal and latitudinal variations with local dominance of super-volatiles

have been recorded by the Rosetta experiment ROSINA. Measurable molecules are also oxygen (O_2), nitrogen (N_2), hydrogen cyanide (HCN) and other compounds of carbon such as methane (CH_4), ethane (C_2H_6), methanal (H_2CO) and methanol (CH_3OH).

Within cometary nuclei volatiles can be in solid forms as well as trapped within amorphous or crystalline water ice. The solid forms derive by condensation from the Solar Nebula at different distances from the proto-Sun being H_2O snow line around 5 AU and CO snow line at around 12 AU. Hence the different content of volatiles can reflect the location in the Solar Nebula at which the cometary grains of a given nucleus might have formed. However, during the cooling of the Solar Nebula, super-volatiles can be also trapped in amorphous or crystalline water ice (in the form of clathrates). In these latter cases super-volatiles are released at much higher temperatures than those of their sublimation.

13.2.4.2 Refractories

The nucleus of 67P/CG appears to be widely coated with dark refractory materials, and water ice have been detected only on freshly exposed surfaces and active sources. The refractory materials have been studied with the Rosetta *Visual InfraRed and Thermal Imaging Spectrometer (VIRTIS)* and mass spectrometers on board the Philae lander, which collected materials lifted up at its first touch-down. They are a complex mixture of organic compounds including aromatic and aliphatic C-H bonds and nitrogen bearing species. Of particular interest are:

- hydroxyethanal (CH_2OHCHO), which is an efficient initiator in the prebiotic formation of sugars,
- methanenitrile (HCN), which is essential in the prebiotic synthesis of amino acids and nucleobases and
- methanamide $HCONH_2$ which provides a prebiotic route to nucleobases.

All the detected compounds can be formed by galactic and solar cosmic rays through ultraviolet (UV) photons or energetic particles irradiation of ices or by the polymerization of mixtures of ices at low temperatures. This would imply that the super-volatile ices (CH_4, CO, CO_2, CH_3OH, etc.), along with water, were readily available at the time of the cometary nucleus formation.

The organic refractories, that place comets among the most favorable candidates for delivering organic materials into the Inner Solar System, are not the unique components of cometary dusts. Indeed, ground-based observations of Hale-Bopp and 9P/Tempel 1 as well as the Stardust samples collected from comet 81P/Wild 2 revealed also the presence of crystalline and amorphous silicates, such as olivine and pyroxene. At the moment of their formation these minerals, characterized by high condensation temperatures (1200–1400 K), were presumably close to Sun and were afterward transported far beyond the snow line to be agglomerated on cometary nuclei. This would imply a strong radial mixing within the Solar Nebula.

13.2.4.3 Isotopic Ratios

One of the major question in Solar System science is from where water and organic compounds on Earth and other terrestrial planets originated. Isotopic ratios might help answering such a question. The D/H ratios show a great variability among the Solar System bodies being low for the Jupiter atmosphere ($2.1 \pm 0.5 \times 10^{-5}$), which is similar to the interstellar ratio and thought to be representative of the proto-solar value, enriched in the inner Solar System bodies (six times the protosolar value) and very high for the long period comets (10–20 times the protosolar value). Hence a general radial trend of enrichment could have characterized the Solar Nebula from lower values close to the Proto-Sun to higher values far away. Exceptions to this rule are the Jupiter family comets showing a great heterogeneity within the same group. Indeed 67P/CG values (5.3×10^{-4}) are consistent to the radial increase, whereas 103P/Hartley 2 and 45P/Honda-Mrkos-Pajdušáková values (1.61×10^{-4} and 2.0×10^{-4}, respectively) are similar to the Earth D/H ratio. For these reasons the Jupiter family comet contribution to Earth oceans and atmosphere is still under debate.

13.2.5 Cometary Geology

Unlike their molecular and atomic composition, which can readily match the primordial accreting material, the structure of cometary nuclei can either reflect the early accretionary processes within the Solar Nebula and Proto-planetary disk or be affected, at various degree, by time-varying collisional environments in the early Solar System, space weathering and intense sublimation, particularly when they move inside the snow-line at a heliocentric distance of about 2.5–3 AU. Hence in the following subsections we have subdivided the cometary nucleus features in (i) primary structures and (ii) erosional morphologies and deposits. In particular, the geological features that might possibly testify the primordial aggregation of the nucleus and all the structures induced by processes other than sublimation and erosion (e.g rotational torque, collisions and thermal fatigue), are here considered primary structures because they constitute the geo-structural skeleton over which cyclic erosional processes take place. All the other geological forms, which are thought to be strictly linked to sublimation and gravitational processes, are here referred as erosional morphologies and deposits. Most of the consideration on cometary nucleus geology of the following subsections are mainly based on observations of comet 67P/CG which have been acquired at an imaging resolution of up to 0.1 m per pixel, greatly exceeding any other previous close-range images, the best of which are around 7 m per pixel.

13.2.5.1 Nuclei Primary Structures

Generally the inner structure of a cometary nucleus is thought to be a chaotic agglomerate of primordial pebbles and fragments (i.e. rubble pile) accreted within the Solar Nebula or welded after a catastrophic collision of a larger parent body. However the geological observations from Rosetta OSIRIS data of the bilobate 67P/CG comet nucleus have highlighted that its skeleton is more ordered than previously supposed and characterized by a thick layered sequence, locally up to 650 m (Fig. 13.12a–d).

This discovery strongly supports the hypothesis that the bi-lobe shape of 67P comet is the result of a merging between two fully formed kilometre-sized cometesimals in the early stages of the Solar System via a low-velocity impact. In particular, the two lobes are strikingly similar in terms of deep onion-like structure (Fig. 13.12a, b), surface composition provided by Rosetta VIRTIS (see Chap. 12) and observed surface features. At the same time, investigations on OSIRIS and Radio Science data have revealed different densities between the two lobes, further supporting their independent origin. Taken together, these structural, physical and compositional features indicate that 67P/CG is most probably composed by two cometesimals, which experienced similar accretion processes, formed independently and, finally, gently fused together. These considerations can be easily extended to comets with similar bi-lobe shape such as 1P/Halley, 19P/Borrelly (Fig. 13.13a), 103P/Hartley or apparent layering such as 81P/Wild 2 , 9P/Tempel 1 (Fig. 13.13b) and require a dynamically colder primordial disk that allows such objects to avoid collisions and survive the age of the Solar System.

As summarised in Chap. 2 a *stratum*, or bed, is the result of a sedimentary process which includes the production, dispersal in a given medium, and deposition of elemental grains. Hence, on the bases of the overall definition of a sedimentary process, layers on 67P/CG and other cometary nuclei can be fully considered as strata. In this view strata on cometary nuclei are by all means the result of the first sedimentary process that ever happened in our Solar System in which grains are formed of dirty icy pebbles and the transport medium is most probably the primordial Solar Nebula (Chap. 9).

The morphological evidences of stratification on cometary nuclei include layered cliffs and terraced walls, cuestas, mesas and hogback morphologies (Fig. 13.14). At high resolution the stratification appears to be locally constituted by metric pebbles forming *Goosebumps* textures. Those might be the expression of the elemental building blocks of primordial cometesimals.

Goosebumps textures can be also the result of pervasive polygonal fractures that, enlarged by focused sublimation, might isolate meter-sized intact blocks. Different kinds of fracture-generating processes such as thermal fatigue, collisional events, rotational torques are indeed all possible on a cometary nucleus (Fig. 13.14a–d). Although almost all fractures are mainly tensile type structures (Mode I), most of them can be attributed to some specific processes on the bases of their geometrical characteristic and location on the nucleus. Thermal cycles modulated by spin periods at perihelion can be extremely effective on a comet. In the case of comet

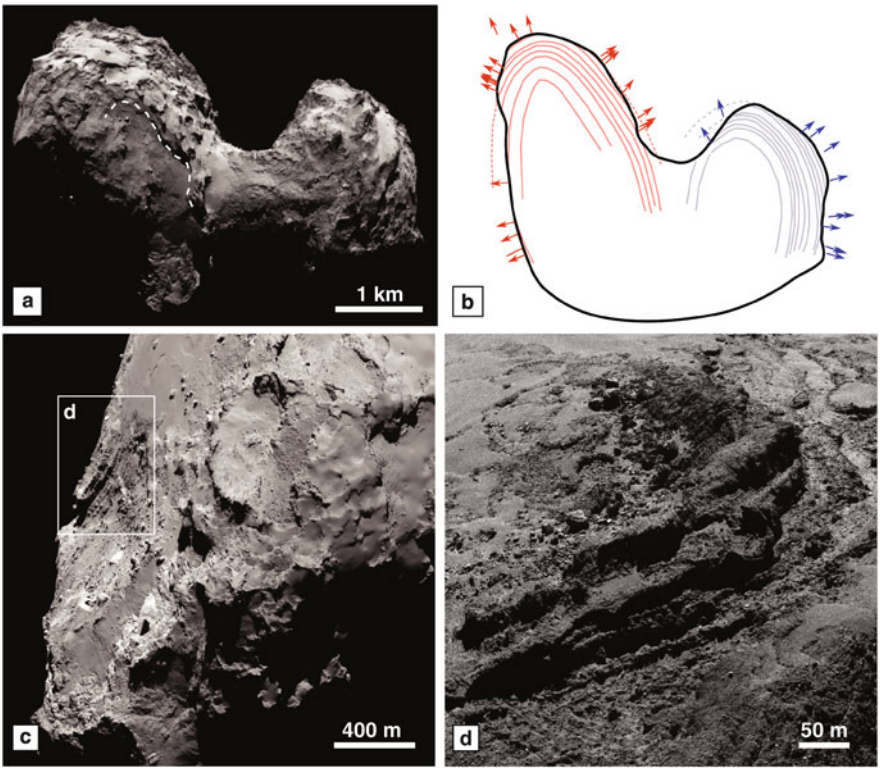

Fig. 13.12 Strata on comet 67P/CG; (**a**) 67P/Churyumov-Gerasimenko (67P/CG) nucleus as seen by Rosetta-OSIRIS Wide Angle Camera on 9 September 2014; note the bi-lobe shape made up of a major lobe (*the body*), a minor one (*the head*) and a neck in-between; the *dashed white line* underlines a strata partially enveloping the body; (**b**) geological section of 67P/CG comet with the interpreted inner stratification; *arrows* are vector perpendicular to strata and terraces on the cometary nucleus (see (**c**) and (**d**)); *red lines* mark strata on the major lobe; *blue lines* mark strata of the minor lobe; the two lobes are independent and characterized by their own onion-like stratification; (**c**) view of a portion of the 67P/CG nucleus acquired by the Rosetta-OSIRIS Narrow Angle Camera on 17 March 2015; note a mesa underlined by a stratification dipping underneath smooth deposits; all around are terraces and cuestas morphologies often covered by dust; the *white square* is the location of (**d**). (**d**) details of the stratification underneath the mesa morphology in (**c**) as imaged by the Rosetta-OSIRIS Narrow Angle Camera image on 19 March 2016. Source: (**a**, **c**, **d**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA. (**b**) redrawn after Massironi et al. 2015. Nature, 526, doi: 10.1038/nature15511

67P/CG, which has a perihelion at 1.3 AU, the temperatures variations might exceed 200 K and the diurnal temperature ranges can reach 15 K/min.

Seasonal thermal weathering can not even be excluded. Any thermal fatigue process creates fractures characteristically arranged in polygonal geometries that in the case of 67P/CG isolate intact blocks with diameters ranging between few to tens of meters (Fig. 13.14a). The thermal fatigue can be so intense that even boulders,

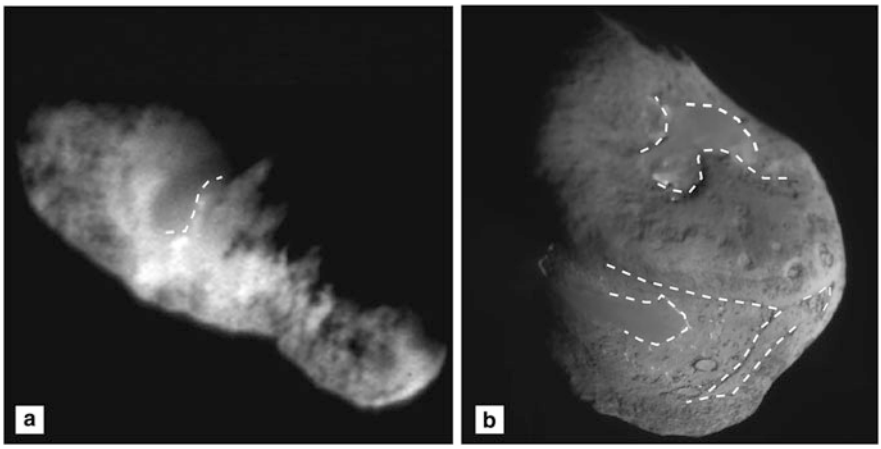

Fig. 13.13 Strata on comets; (**a**) 19P/Borrelly bi-lobe comet as seen by Deep Space 1 (DS1) spacecraft in September 2001; note the smooth region bordered by a terrace margin (*dashed white line*), which can be an evidence of layering; (**b**) 9P/Tempel 1 comet acquired by Deep Impact spacecraft; note terrace margins (*dashed lines*) that suggest layering and roundish depressions similar to the one found on Wild 2 (Fig. 13.16) and 67P/CG. Source: (**a**) NASA Planetary Photojournal. (**b**) NASA/JPL/UMD

detached from cliffs and walls, can be affected by a thermally induced intimate fracturing (Fig. 13.14b).

The gravitational pool of the sun or large planetary bodies onto cometary nuclei as well as the sublimation jets activated during perihelion passages might lead to excited rotation states and induce the opening of straight and conjugate tensile fractures with potential minor strike-slip components. A good example is the 67P/CG comet nucleus where the rotational axis crosses a neck region separating two lobes of different size and possibly slightly different density. In such a case it has been demonstrated that inner stresses, induced by the rotational torques, are focused on the neck region and led to the opening of 200 m long fractures (Fig. 13.14c).

Finally systems of sharp fracture planes can be derived by collisional events related to primordial cometesimal aggregation or later impacts. Indeed, it has been proposed that the noticeable fracture system transecting the stratification on the most spectacular wall of 67P/CG comet (the 900 m high Hathor cliff) has been originated by the original margining of the two lobes (Fig. 13.14d).

Hence finding any primordial structures on a cometary nucleus is not a trivial task since cometary surfaces are interested by numerous fracture-generating events as well as solar energy-activated processes that can deeply rework and mask primary structures.

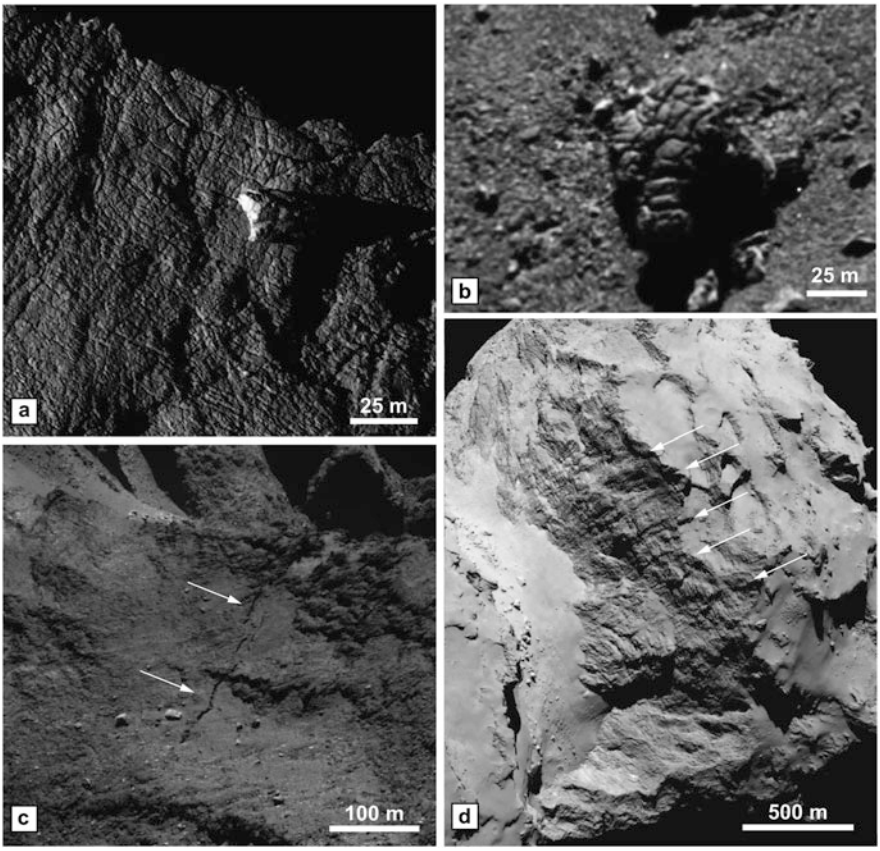

Fig. 13.14 Fractures on 67P/CG; (**a**) polygonal fractures on 67P/CG comet due to thermal fatigue; (**b**) fractured boulder on 67P/CG; (**c**) 500 m long fracture at the 67P/CG neck region induced by rotational torque; (**d**) Parallel lineaments crosscutting a layered sequence on a 900 m high wall of the 67P/CG head. Source: (**a**–**d**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/-SSO/INTA/UPM/DASP/IDA

13.2.5.2 Nuclei Erosional Morphologies and Deposits

A cometary nucleus is the realm of sublimation given the high quantity of mass that any nucleus loses at its perihelion passage. For example the dust loss at perihelion for a short-period comet of small size such as 67P/CG (4×2 km) was around 1000 kg/s whereas long-period comets such as Hyakutake (3 km of diameter) and Hale-Bopp (40 km of diameter) had a loss rate of 10,000 and 400,000 kg/s respectively.

The differential outgassing activity on cometary nuclei can reflect an uneven distribution of refractory dust covers or even inner compositional heterogeneities. Indeed the surface of a cometary nucleus is characterized by smooth areas covered by dusty lag material deposited after sublimation and outcrops of rocky-like

appearance. Sharp retreating scarps, active pits, roundish depressions, thin surface stratification (due to subsurface water vapour rise and crystallization) are large-scale examples of features related to cometary activity; *honey-combs*, aligned pits and aeolian-like morphologies are the most prominent examples of small-scale features.

Scarps underlining terraces and mesas are common and renown features on nucleus surfaces being discovered on 19P/Borrelly, 81P/Wild 2, and 9P/Tempel 1 and classically attributed to slope retreats driven by sublimation. This process is very effective on comet 67P/CG where sublimation most probably focuses along thermally induced fractures. The differential erosion across the stratification on comet 67P/CG leads to cuestas morphologies and points in favor of a certain compositional heterogeneity among strata (Fig. 13.14a).

Scarp retreats are intimately associated to rock-fall type gravitational phenomena leading to landslide accumulations and taluses (Chap. 9). Boulders of such deposits are characterized by size-frequency distributions giving hints on their age. In particular freshly collapsed deposits, still along the scarps or at their bases, are characterized by power law indices ranging from −6 to −4, whereas isolated ancient deposits, which are far from the surrounding cliffs, have lower indices of around −1 to −2 (Fig. 13.15a, b) These results are most probably due to the concurrent effects of dust mantling and erosional sublimation of small boulders of older deposits.

When Stardust spacecraft approached 81P/Wild 2 nucleus in 2004, it was clear how a cometary surface can be extremely variegated in terms of morphologies and classically dominated by roundish features (Fig. 13.16a–d). The roundish depressions, afterward well seen also in 9P/Tempel 1, are characterized by rim-less morphologies and steep to sub-vertical slopes. They were firstly explained either

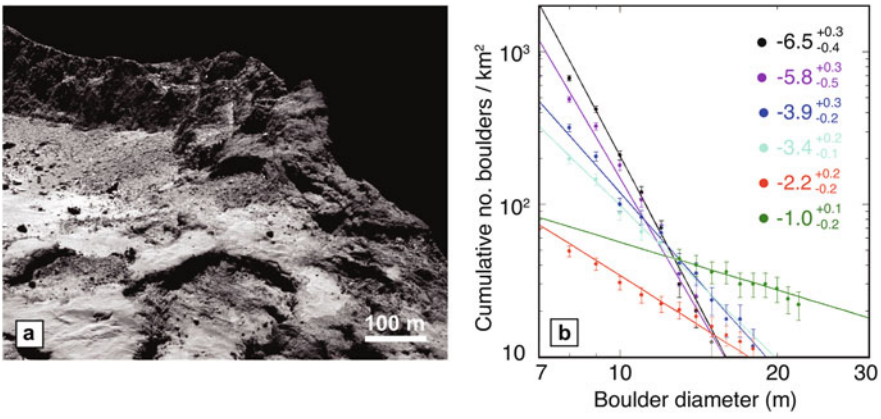

Fig. 13.15 Features on 67P/CG; (**a**) gravitational deposits at a cliff foot on 67P/CG; (**b**) boulder size-frequency distribution of various deposits on 67P/CG; steeper distributions are younger gravitational deposits induced by sublimation, shallower distributions are mature deposits which most probably underwent a prolonged sublimation activity. Source: (**a**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA. (**b**) Redrawn after Pajola et al. 2015. Astronomy & Astrophysics, 583, doi: 10.1051/0004-6361/201525975

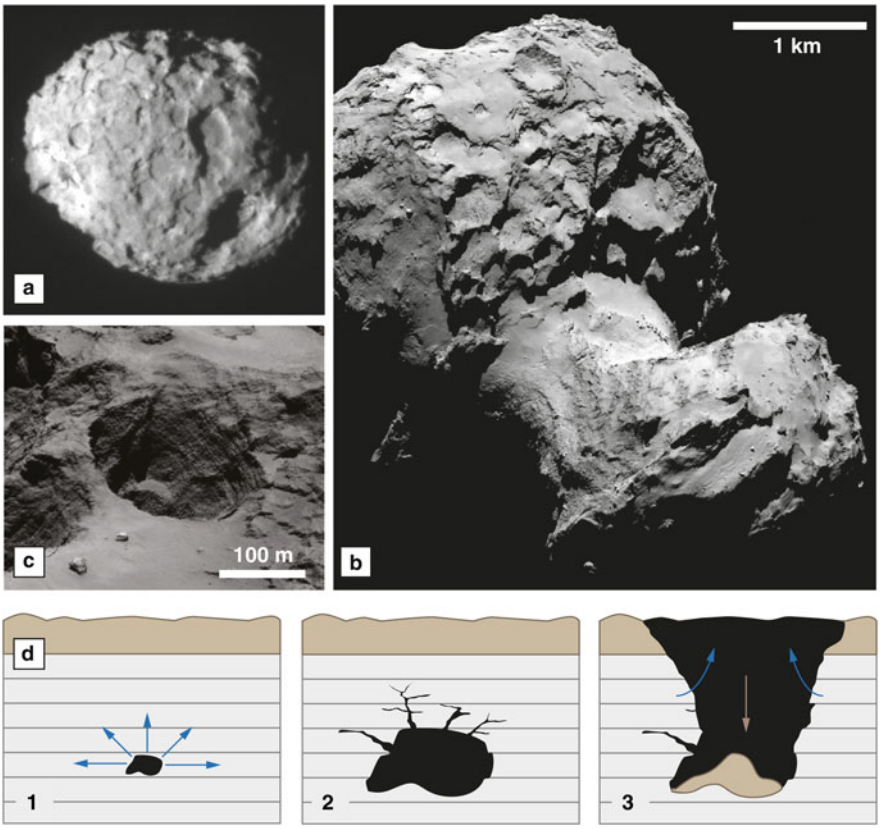

Fig. 13.16 Roundish depressions on cometary surfaces; (**a**) Wild 2 comet acquired by NASA Stardust spacecraft on 2 January 2004. Note the widespread roundish depressions over the surface; (**b**) 67P/CG northern hemisphere as seen by Rosetta-OSIRIS NAC camera in August 2014; note the numerous roundish terraces and depression on the larger lobe; (**c**) a 200 m wide and 20 m deep roundish pit on 67P/CG; (**d**) proposed process of generation of roundish pits: a cavity forms (1) and expands (2) due to subsurface heat and sublimation; the gas can reach the surface trough fractures whereas the cavity expands until the roof collapse (3). Source: (**a**) NASA/JPL/STARDUST. (**b**, **c**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA. (**d**) Modified after Vincent et al. 2015. Nature, 523, doi: 10.1038/nature14564

as impact tunnels on soft materials (followed by strong sublimation processes) or as generated by outburst activity. The roundish mesas were covertly described as remnants of ancient gas conduits consolidated by crystallized water ice particles and exposed by differential erosion. However with the Rosetta mission to comet 67P/CG it was evident that pristine crater morphologies are very rare on a cometary surface because of the intense erosional activity due to sublimation and the deepest roundish morphologies are clear centre of jet emissions (Figs. 13.16a–d and 13.17a, b). These forms (up to 200 m deep and with a depth/diameter ratio close to 1) were then called *active pits* and their genesis is explained as a concurrent process of internal erosion,

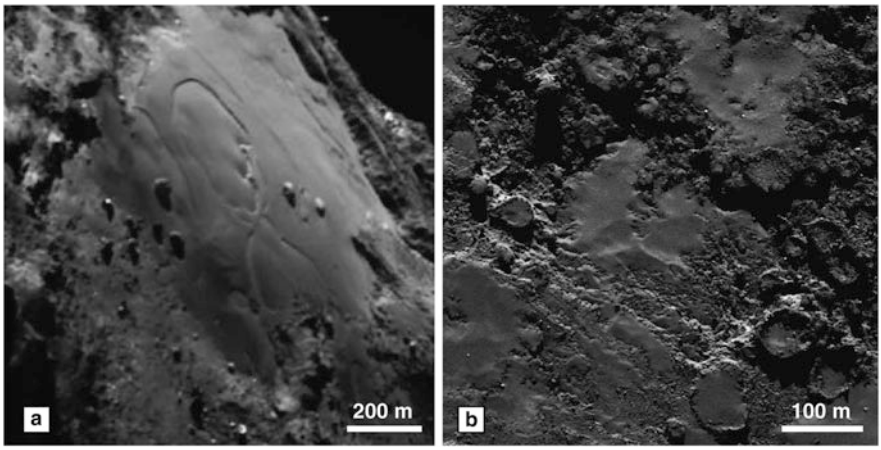

Fig. 13.17 Roundish depressions on cometary surfaces; (**a**) two roundish depression on comet 67P/CG. These features appeared as small depression in the first half of June 2015 and expanded at a rate of $5–8\times10^{-5}$ m/s until they reached a diameter of around 200 m on the 2nd of July 2015 when this image was taken by the Rosetta/OSIRIS-NAC camera; (**b**) roundish mesas on comet 67P/CG. Source: ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

gravitational failures and jet emissions. In particular, it is thought that sublimation of super-volatile materials at depth and their escape through percolating fracture systems leads to subsurface voids. The heat required for subsurface sublimation is thought to be derived laterally from the exposed scarps or due to energetic processes such as clathrate destabilization or amorphous water ice crystallization. The expansion of subsurface cavities and the gas escape might trigger outbursts and cause collapses of the cavity roofs exposing new fresh walls, free of dust covers. On such inner walls sublimation focuses generating collimated jets visible on the Rosetta images.

The active pits can than evolve by wall retreat and partial gravitational collapse, giving rise to spectacular landscapes made up of sequences of cliffs and roundish denudation terraces (Fig. 13.16b). Peculiar flow-like deposits have been detected associated to one of the dormant active pits on 67P/CG. Indeed foams inflated by sublimating gasses might expand from emission centres and give rise to these overlapping effusions. This process has been proposed to explain also some smooth terrains limited by fan-shaped and lobate margins on Tempel 1.

Other large-scale roundish depressions have been seen during their development when comet 67P/ approached perihelion in 2015. They are up to 200 m large and expanded at a rate of $5–8\times10^{-5}$ m/s, exposing few meters high scarps at their margins. Their genesis has been related to surface erosion by sublimation, but internal bulges within the depressions might also suggest possible endogenic processes like the ones responsible for active pits formation (Fig. 13.18a).

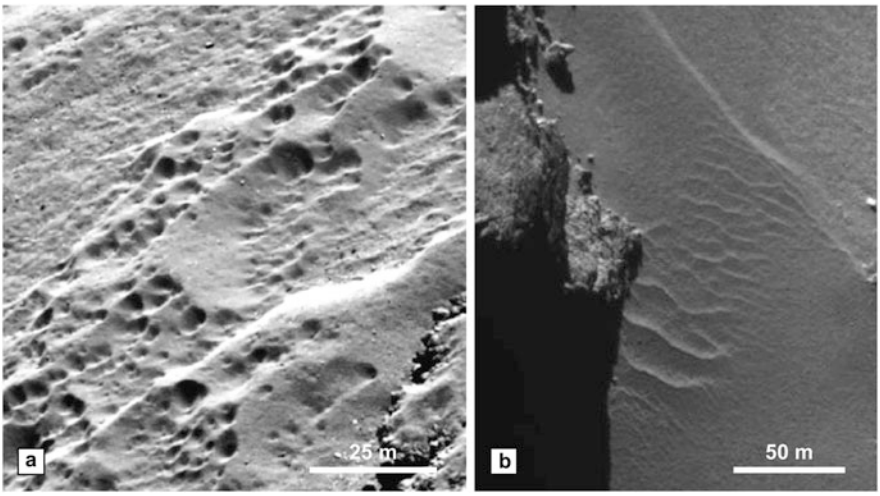

Fig. 13.18 Surface features on 67P/CG; (**a**) meter-size pits aligned along sun-facing slopes on 67P/CG comet; (**b**) dune-like morphologies on 67P/CG. Source: (**a**, **b**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

The roundish mesas have not found yet a convincing explanation, they might be remnants of ancient conduits or simply the result of slope retreat on a compositionally inhomogeneous material (Fig. 13.18b).

Sublimation is not only concentrated in specific locations along scarps or active pits but can be widespread on large areas like in the case of the neck junction between the two lobes of 67P/CG comet, made up of slightly bluer loose materials which started sublimating in May 2014 at more than 3.5 AU from the Sun. However the areal sublimation events, recorded on 67P/CG and 9P/Tempel 1, have shown that this kind of activity is only partially correlated to exposed water ice. On 67P/CG was, in particular, demonstrated that water ice appears only temporally and cyclically on such locations, in function of the local sun illumination conditions during the day. This effect has been explained by water day cycles where water vapor percolates through fractures and pores from deeper layers to the surface during the day, crystallizes as water ice close to the surface during the night and is suddenly released in the early morning. This daily process affects only the uppermost centimeters of the cometary nucleus, but water vapor ascent and crystallization over several orbits might even create a hardened water-ice rich zone which can extend up to 10 m underneath the refractory dust cover. This process does not substitute the original stratigraphy made up of 100 m thick sequences of layered material, but it likely affects the rheological properties of the uppermost part of the nucleus which should be hardened though sintering or grain growth by vapor diffusion.

Among the small forms related to sublimation, pitted terrains called *honeycombs* are the most prominent ones. They are meter-size features highly resembling morphologies due to differential sublimation on other planetary bodies (Chap. 11)

such as *hollows* on Mercury and *swiss cheese* on Mars. These morphologies form on dusty areas and seem to be strictly related to sunset jets which are triggered by water ice sublimation and sustained by thermal lag in the dusty subsurface up to 2 h beyond the sunset. Meter-size pits aligned along sun-facing slopes at local midday (Fig. 13.17) might have similar genesis although a direct observation of their formation has not been done yet.

Finally several aeolian-like morphologies such as ripples, wind tails behind boulders and putative dunes have been detected on 67P/CG comet nucleus (Fig. 13.18b). The mechanism responsible for the formation of these features is still unknown, although surface dust transport by gas dragging is one of the most plausible candidates. However in this case emission centres should be very close to the aeolian-like forms since as the gas expands the pressure profile rapidly decreases preventing gas drag to overcome the grain cohesion, which is the dominant force preventing dust particle motion in such a low gravity environment. An alternative and more effective process is the air-fall splash effect. It has been demonstrated that the dust mantle covering a cometary nucleus can be originated either by lag materials left in site after sublimation or air-fall particles that where formerly lifted by sublimation and, failing to escape the nucleus, re-impacted its surface. Air-fall impacts (splashes) have enough energy to overcome particle cohesion and mobilize grains. This makes the needed velocities of gas streams for creating aeolian-like morphologies much lower than the threshold frictional speed. This process seems the best candidate for explaining ripples and wind tail formation, but is not able to account for the larger putative dunes. Indeed the latter ones are always aligned to the sublimating pit chains, thus they can be the result of material transport and lateral storage from the gas stream emitted by these small sublimating centres.

13.2.5.3 Geomechanical Properties

The mechanical properties of the cometary nucleus materials are quite debated. The compressive strength of the uppermost consolidated layer, potentially affected by intense solar processing, has been estimated from the Philae bouncing at its first touch-down to be in the order of a few kPa. However the penetrometric measurements made at Philae landing site gave values of ca 2 MPa, similar to that of sintered porous ice. The different scale of measurement being metric in the first case and centimetric in the second could affect the estimates. This indeed means that the penetrometer was measuring a homogeneous material whereas the bouncing estimates are instead related to a larger mass, which might include possible heterogeneities. Because of comets' high porosity, many authors believe that these values are not representative of the pristine material which should be much more *fluffy*. Indeed estimates based on collapsed overhangs tens to hundreds meters large gave values of tensile strength from few Pa to less than 150 Pa (for comparison, 100 kg m^{-3} of snow has a tensile strength of 100 Pa). According to these values the compressive strength should range between hundreds Pa to less than 1.5 kPa. However, also in this case the values are related to a mass affected by pervasive discontinuities due to fracturing.

Take-Home Messages

Most asteroids are located in the asteroid Main Belt between Mars and Jupiter.
Most meteorites sample asteroidal bodies.
The knowledge of asteroidal composition is based on observed spectral properties and analyzed meteorites.
Asteroids are dynamically influenced by the major planets and able to leave the Main Belt.
The geologic evolution of asteroids is characterized by early radioactive heating (differentiated bodies) and subsequent collision and impact gardening.
Comets formed beyond the snow line of the proto-planetary disk and were stored in the cold outer Solar System reservoirs, remaining potentially unaffected by major modifications for 4.5 Gyr until to their deflection into the present orbits and consequent activation.
Comets are fossils of the nascent Solar System with molecular and atomic composition matching the primordial accreting material and nuclei structures possibly still reflecting the early accretionary processes (e.g cometary deep stratification).
Cometary nuclei are the realm of sublimation processes with peculiar landforms such as active pits, honeycombs, roundish active depressions, retreating scarps and denudation terraces.
Cometary material can be fractured by numerous processes such as thermal fatigue, collisional events and rotational torque induced by outgasing events, tidal stress or the change of the center of mass due to unbalanced loss of mass.
Sedimentary transport and deposition processes on comets are possible though airfall of particles formerly lifted by sublimation and gas dragging.

Suggested Readings

Bottke, W.F., Cellino, A., Paolicchi, P., Binzel, R.P. (eds.): Asteroids III. Space Science Series. University of Arizona Press, Tucson (2002)

Bottke, W.F., Nolan, M.C., Greenberg, R., Kolvoord, R.A.: Velocity distributions among colliding asteroids. Icarus **107**, 255–268 (1994). doi:10.1006/icar.1994.1021

Capaccioni, F., Coradini, A., Filacchione, G., Erard, S., Arnold, G., Drossart, P., De Sanctis, M.C., Bockelee-Morvan, D., Capria, M.T., Tosi, F., Leyrat, C., Schmitt, B., Quirico, E., Cerroni, P., Mennella, V., Raponi, A., Ciarniello, M., McCord, T., Moroz, L., Palomba, E., Ammannito, E., Barucci, M.A., Bellucci, G., Benkhoff, J., Bibring, J.P., Blanco, A., Blecka, M., Carlson, R., Carsenty, U., Colangeli, L., Combes, M., Combi, M., Crovisier, J., Encrenaz, T., Federico, C., Fink, U., Fonti, S., Ip, W.H., Irwin, P., Jaumann, R., Kuehrt, E., Langevin, Y., Magni, G., Mottola, S., Orofino, V., Palumbo, P., Piccioni, G., Schade, U., Taylor, F., Tiphene, D., Tozzi, G.P., Beck, P., Biver, N., Bonal, L., Combe, J.-P., Despan, D., Flamini, E., Fornasier, S., Frigeri, A., Grassi, D., Gudipati, M., Longobardo, A., Markus, K., Merlin, F., Orosei, R., Rinaldi, G., Stephan, K., Cartacci, M., Cicchetti, A., Giuppi, S., Hello, Y., Henry, F., Jacquinod, S., Noschese, R., Peter, G., Politi, R., Reess, J.M., Semery, A.: The organic-rich surface of comet 67P/Churyumov-Gerasimenko as seen by VIRTIS/Rosetta. Science **347**(6220), 0628-1–0628-4 (2015). doi:10.1126/science.aaa0628

Hässig, M., Altwegg, K., Balsiger, H., Bar-Nun, A., Berthelier, J. J., Bieler, A., Bochsler, P., Briois, C., Calmonte, U., Combi, M., De Keyser, J., Eberhardt, P., Fiethe, B., Fuselier, S.A., Galand, M., Gasc, S., Gombosi, T.I., Hansen, K.C., Jäckel, A., Keller, H.U., Kopp, E., Korth, A., Kührt, E., Le Roy, L., Mall, U., Marty, B., Mousis, O., Neefs, E., Owen, T., Rème, H., Rubin, M., Sémon, T., Tornow, C., Tzou, C.-Y., Waite, J.H., Wurz, P.: Time variability and heterogeneity in the coma of 67P/Churyumov-Gerasimenko. Science **347**(6220), 0276-1–0276-4 (2015). doi:10.1126/science.aaa0276

Massironi, M., Simioni, E., Marzari, F., Cremonese, G., Giacomini, L., Pajola, M., Jorda, L., Naletto, G., Lowry, S., El-Maarry, M. R., Preusker, F., Scholten, F., Sierks, H., Barbieri, C., Lamy, P., Rodrigo, R., Koschny, D., Rickman, H., Keller, H.U., A'Hearn, M. F., Agarwal, J., Auger, A.-T., Barucci, M.A., Bertaux, J.-L., Bertini, I., Besse, S., Bodewits, D., Capanna, C., da Deppo, V., Davidsson, B., Debei, S., de Cecco, M., Ferri, F., Fornasier, S., Fulle, M., Gaskell, R., Groussin, O., Gutiérrez, P.J., Güttler, C., Hviid, S. F., Ip, W.-H., Knollenberg, J., Kovacs, G., Kramm, R., Kührt, E., Küppers, M., La Forgia, F., Lara, L.M., Lazzarin, M., Lin, Z.-Y., Lopez Moreno, J.J., Magrin, S., Michalik, H., Mottola, S., Oklay, N., Pommerol, A., Thomas, N., Tubiana, C., Vincent, J.-B.: Two independent and primitive envelopes of the bilobate nucleus of comet 67P. Nature, **526**, 402–405 (2015). doi:10.1038/nature15511

Pajola, M., Vincent, J.-B., Güttler, C., Lee, J.-C., Bertini, I., Massironi, M., Simioni, E., Marzari, F., Giacomini, L., Lucchetti, A., Barbieri, C., Cremonese, G., Naletto, G., Pommerol, A., El-Maarry, M. R., Besse, S., Küppers, M., La Forgia, F., Lazzarin, M., Thomas, N., Auger, A.-T., Sierks, H., Lamy, P., Rodrigo, R., Koschny, D., Rickman, H., Keller, H.U., Agarwal, J., A'Hearn, M.F., Barucci, M.A., Bertaux, J.-L., Da Deppo, V., Davidsson, B., De Cecco, M., Debei, S., Ferri, F., Fornasier, S., Fulle, M., Groussin, O., Gutierrez, P.J., Hviid, S.F., Ip, W.-H., Jorda, L., Knollenberg, J., Kramm, J.-R., Kürt, E., Lara, L.M., Lin, Z.-Y., Lopez Moreno, J.J., Magrin, S., Marchi, S., Michalik, H., Moissl, R., Mottola, S., Oklay, N., Preusker, F., Scholten, F., Tubiana, C.: Size-frequency distribution of boulders $\geq$7 m on comet 67P/Churyumov-Gerasimenko. Astron. Astrophys. **583**, A37 (2015). doi:10.1051/0004-6361/201525975

Prettyman, T.H., Mittlefehldt, D.W., Yamashita, N., Lawrence, D.J., Beck, A.W., Feldman, W.C., McCoy, T.J., McSween, H.Y., Toplis, M.J., Titus, T.N., Tricarico, P., Reedy, R.C., Hendricks, J.S., Forni, O., Le Corre, L., Li, J.-Y., Mizzon, H., Reddy, V., Raymond, C.A., Russell, C.T.: Elemental mapping by Dawn reveals exogenic H in Vesta's regolith. Science **338**, 242 (2012). doi:10.1126/science.1225354

Tang, H., Dauphas, N.: Abundance, distribution, and origin of $^{6}{}_{0}$Fe in the solar protoplanetary disk. Earth Planet. Sci. Lett. **359**, 248–263 (2012). doi:10.1016/j.epsl.2012.10.011

Tera, F., Papanastassiou, D., Wasserburg, G.: The lunar time scale and a summary of isotopic evidence for a terminal lunar cataclysm. Lunar Planet. Sci. Conf. Abstr. **5**, 792 (1974)

Thomas, N., Sierks, H., Barbieri, C., Lamy, P.L., Rodrigo, R., Rickman, H., Koschny, D., Keller, H.U., Agarwal, J., A'Hearn, M.F., Angrilli, F., Auger, A.-T., Barucci, M.A., Bertaux, J.-L., Bertini, I., Besse, S., Bodewits, D., Cremonese, G., Da Deppo, V., Davidsson, B., De Cecco, M., Debei, S., El-Maarry, M.R., Ferri, F., Fornasier, S., Fulle, M., Giacomini, L., Groussin, O., Gutierrez, P.J., Güttler, C., Hviid, S.F., Ip, W.-H., Jorda, L., Knollenberg, J., Kramm, J.-R., Kührt, E., Küppers, M., La Forgia, F., Lara, L.M., Lazzarin, M., Moreno, J.J.L., Magrin, S., Marchi, S., Marzari, F., Massironi, M., Michalik, H., Moissl, R., Mottola, S., Naletto, G., Oklay, N., Pajola, M., Pommerol, A., Preusker, F., Sabau, L., Scholten, F., Snodgrass, C., Tubiana, C., Vincent, J.-B., Wenzel, K.-P.: The morphological diversity of comet 67P/Churyumov-Gerasimenko. Science **347**(6220), 0440-1–0440-6 (2015). doi: 10.1126/science.aaa0440

Vincent, J.-B., Bodewits, D., Besse, S., Sierks, H., Barbieri, C., Lamy, P., Rodrigo, R., Koschny, D., Rickman, H., Keller, H.U., Agarwal, J., A'Hearn, M.F., Auger, A.-T., Barucci, M.A., Bertaux, J.-L., Bertini, I., Capanna, C., Cremonese, G., da Deppo, V., Davidsson, B., Debei, S., de Cecco, M., El-Maarry, M.R., Ferri, F., Fornasier, S., Fulle, M., Gaskell, R., Giacomini, L., Groussin, O., Guilbert-Lepoutre, A., Gutierrez-Marques, P., Gutiérrez, P.J., Güttler, C., Hoekzema, N., Höfner, S., Hviid, S.F., Ip, W.-H., Jorda, L., Knollenberg, J., Kovacs, G., Kramm, R., Kührt, E., Küppers, M., La Forgia, F., Lara, L.M., Lazzarin, M., Lee, V., Leyrat, C., Lin, Z.-Y., Lopez Moreno, J.J., Lowry, S., Magrin, S., Maquet, L., Marchi, S., Marzari, F., Massironi, M.,

Michalik, H., Moissl, R., Mottola, S., Naletto, G., Oklay, N., Pajola, M., Preusker, F., Scholten, F., Thomas, N., Toth, I., Tubiana, C.: Large heterogeneities in comet 67P as revealed by active pits from sinkhole collapse. Nature **523**, 63–66 (2015). doi: 10.1038/nature14564

Weidenschilling, S.J.: The distribution of mass in the planetary system and solar nebula. Astrophys. Space Sci. **51**(1), 153–158 (1977). doi:10.1007/BF00642464

Part IV
Frontiers

Chapter 14
Astrobiology, the Emergence of Life, and Planetary Exploration

Barbara Cavalazzi, Mihaela Glamoclija, André Brack, Frances Westall, Roberto Orosei, and Sherry L. Cady

14.1 Astrobiology

Astrobiology is the study of the origin, evolution, and distribution of life in the context of the cosmic evolution in our Solar System and beyond. As such, it crosscuts a number of disciplines that include biology, astronomy, geology, chemistry, physics, engineering, computational science, planetary protection, and philosophy. Astrobiologists seek to understand life's origin, the nature of early life on Earth, and how life might have emerged, adapted, and survived elsewhere in the universe.

B. Cavalazzi (✉)
Università di Bologna, Bologna, Italy
e-mail: barbara.cavalazzi@unibo.it

M. Glamoclija
Rutgers University, New Brunswick, NJ, USA
e-mail: m.glamoclija@rutgers.edu

A. Brack • F. Westall
Centre de Biophysique Moléeculaire, Orléans, France
e-mail: brack@cnrs-orleans.fr; westall@cnrs-orleans.fr

R. Orosei
Istituto Nazionale di Astrofisica, Bologna, Italy
e-mail: roberto.orosei@inaf.it

S. L. Cady
Pacific Northwest National Laboratory, Richland, WA, USA
e-mail: sherry.cady@pnnl.gov

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_14

The term *astrobiology*, which was introduced in 1941 by A. L. LAFLEUR, is often used interchangeably with the word *exobiology*[1] introduced by J. LEDERBERG in 1960. This discipline has also been called *astrobotany* (by G. A. TIKHOV in 1945), *cosmobiology* by J. D. BERNAL (1901–1971) and *bioastronomy* by the *International Astronomical Union Commission* in 1982, though these terms are now obsolete. Although astrobiology is considered to be an emerging discipline, it cannot be considered a new scientific discipline. ARY STERNFELD (1905–1980), a pioneer of astronautics, published incredibly original and modern ideas about astrobiology in 1935, and the first symposium in astrobiology was held in the United States in 1957.

Ancient cultures appear to have been fascinated by the potential for life elsewhere long before the advent of the telescope, the scientific method, or rocketry. More than 400 BC, Greek philosophers, such as THALES OF MILETUS (624–546), EPICURUS (341–270), and METRODORUS OF LAMPSACUS (331–278) mentioned the possibility of the multiplicity of worlds with possible life, and believed that Earth was not the only inhabited world in existence. Centuries later, THOMAS DIGGES (1546–1595) and GIORDANO BRUNO (1548–1600), supporting the Copernican idea that Earth is not at the center of our Solar System, also argued for an infinite number of inhabited worlds in an infinite universe. More than 400 years later, the first exoplanets were discovered, and the prospects for life detection beyond Earth and our Solar System, completely changed.

The age of space investigations began with the work of the first astronomers such as GALILEO GALILEI (1564–1642) JOHANNES KEPLER (1571–1630) GIOVANNI CASSINI (1625–1712) and CHRISTIAAN HUYGENS (1629–1695). Their discovery of other planets and moons revealed that Earth was not central to the universe and that it might not be the only inhabited world in it. The age of space exploration started officially in 1957 with the launch of the first artificial satellite and the first orbiting terrestrial organism (Sputnik Missions). In 1969, the first human walked on the Moon (Apollo 11 Mission).

Today, the frontier of space exploration spans the universe. Advanced instruments onboard the NASA Mars Science Laboratory mission (Curiosity rover) and Mars Exploration Rover mission (with rovers Spirit and Opportunity) have demonstrated the potential for nearby Mars to host extant or extinct life. The NASA Cassini spacecraft images of active geysers on Saturn's moon Enceladus demonstrate the dynamic geophysical nature of planetary bodies in our Solar System. Spectacular three-dimensional images of Comet 67P/Churyumov-Gerasimenko from the European Rosetta spacecraft and flyby images of Pluto from the New Horizon spacecraft exemplify the potential for unexpected discovery.

The continuous monitoring of stars beyond our Solar System by the Hubble Space Telescope and, especially, the discoveries of new star systems by the Kepler

[1]The word exobiology, which is the study of life beyond Earth's atmosphere, is now generally used to refer to studies of the origin and evolution of life, whereas the word astrobiology is used when considering the origin and evolution of planets as a context for life.

Space Telescope illustrate the infinite possibilities for astrobiological exploration. The number of exoplanets orbiting Sun-like stars that has been identified by ground-based and space-based telescopes (e.g., Hubble Space Telescope orbiting Earth since 1990; NASA's Kepler Space Telescope since 2009) has transformed our understanding of planetary habitability. NASA's Kepler mission results, in particular, show that Earth-sized exoplanets orbiting in the habitable zone of small stars are common, thus considerably expand our desire to look for life elsewhere in the universe.

In spite the fact that space is infinitely large with infinite possibilities for exploration and life, the question are we alone in the universe? is a still resistant to an answer. Trying to answer this question, astrobiologists have to face first order question of how to define life.

Perhaps the most general working definition of life is the one adopted by the NASA Exobiology Program in October 1992:

> Life is a self-sustained chemical system capable of undergoing Darwinian evolution.

With this definition for life, which implies that the system uses external matter and energy provided by the environment, primitive life could then be defined as an open chemical system capable of self-reproduction. In other words, primitive life is able to make more of itself by itself and is capable of evolving as a chemical system. During evolution, occasional random errors occur in the transfer of chemical information that may increase the complexity and efficiency of the system and enhance its adaptation to changes within the constraints of its immediate environment. From an astrobiological perspective, any definition of life must be based upon the common characteristics possessed by all Earth life, while considering the possibility that life could be radically different from terrestrial, living systems. Since all living organisms generate (bio)signatures that, if preserved in the geological record, would evidence life existence, therefore astrobiologists must consider a wide range of possible biosignatures that may be indicative of presence of life beyond Earth.

14.2 The Emergence of Life

The origin of terrestrial life remains a working hypothesis, which is why any hypotheses and speculation about the origin of life in science and philosophy must consider the fact that Earth is the only known oasis of life in our Solar System and the universe. The hypothesis that life has a terrestrial origin is suggested by the results of a number of experiments performed in areas of synthetic biology and prebiotic chemistry. An alternative hypothesis is that life originated and proliferated on other planets and was transferred to Earth through space (i.e., the Panspermia hypothesis). The ubiquitous distribution of bio-essential elements (e.g., C, H, N, O, P, S) and life-based molecules and compounds in the cosmos and the variety of possible ways life could be transferred through the universe (e.g., dust, meteorites, asteroids etc.) suggests that life could also have appeared elsewhere on other planets.

In this chapter, we focus only on the origin of terrestrial life and the pioneering prebiotic chemistry phase of this research.

14.2.1 The Chemical Origin of Life

It has been widely accepted that life emerged as a heterotrophic living entity in an aqueous *primordial soup* after JOHN B.S. HALDANE (1892–1964) by chemical reaction of preformed organic molecules (molecules that contain carbon and hydrogen associated with oxygen, nitrogen, and sulfur). Such prebiotic molecules could have formed by a variety of processes that are consistent with conditions on the early Earth.

The simplest sources of carbon in prebiotic organic molecules are gaseous compounds, such as carbon dioxide, and carbon monoxide (oxidized carbon compounds), and methane (a reduced carbon compound). STANLEY L. MILLER (1930–2007) exposed a mixture of methane, ammonia, hydrogen, and water vapor to spark discharge and silent electric discharge, and obtained 2 of the 22 naturally occurring amino acids (glycine and alanine) via the intermediary formation of hydrogen cyanide and aldehydes. Miller's laboratory synthesis of amino acids occurs efficiently when a reducing gas mixture containing significant amounts of hydrogen is used. Today, however, the dominant view is that the primitive atmosphere consisted mainly of CO_2, N_2, and H_2O, along with small amounts of CO and H_2. When Miller's laboratory experiment is run with such a mixture, it yields only small amounts of amino acids, including aspartic and glutamic acids, serine, glycine, alanine, β-alanine, α-aminobutyric acid, and γ-aminoisobutyric acid, obtained when oxidation inhibitors, such as ferrous iron, are added to the system.

The reducing conditions in hydrothermal systems were also an important source of endogenous building blocks of life on the primitive Earth. When the reduced compounds brought to the surface in hydrothermal systems mix with cold (2–4 °C) ocean water, inorganic sulfides precipitate. Though hydrothermal vents are often disqualified as efficient reactors for the synthesis of prebiotic organic molecules because of the high temperatures of their fluids, the organic products that are synthesized in hot vents are rapidly quenched by the surrounding cold water, which may preserve them.

A large fraction of organic matter on primitive Earth was of extraterrestrial origin, as documented by the presence of carbonaceous components in meteorites. Carbonaceous chondrites contain from 1.5–4 wt% carbon (Chap. 6), for the most part as organic compounds. The life cycle of interstellar amino acids, which form in the interstellar medium and land on Earth in meteorites, has been tested in the laboratory and in space. Micrometeorites collected in Antarctica indicate that about 3×10^{19} kg of complex carbon molecules were delivered to primitive Earth over 200 Myr.

By analogy with contemporary living systems, it is tempting to postulate that primitive life may have emerged as a cell-like system that required at a minimum the combination of pre-RNA molecules capable of storing and transferring the information needed for reproduction, pre-enzymes that would have performed the basic chemical work, and pre-membranes able to isolate the system from the aqueous environment. Since RNA can act both as an information molecule and a catalytic molecule, it has been considered as a candidate for the first living system that preceded the cellular world. The structural units that make up RNA are complicated, however, and alternative scenarios have been proposed. The spontaneous organization of amphiphilic molecules to form vesicles has been postulated as the first step toward the origin of life. The primordial role of lipid vesicles as proto-cells has been recently documented. Chemists also consider the potential that primitive self-replicating systems depended on simple autocatalytic molecules adsorbed on solid surfaces, which could solve some of the problems of high dilution of organics in the primitive oceans. In opposition to the idea of a *primordial soup* generating chemical automata from pre-formed building blocks, it has been proposed that primitive autotrophic living entities obtained their ingredients from very simple substances present in their surroundings. Carbon dioxide was abundant in the primitive atmosphere. The energy source needed for reduction of carbon dioxide may have been derived from the oxidation of iron sulfide and hydrogen sulfide during pyrite formation. Pyrite has positive surface charges and will bond the products of carbon dioxide reduction, making a two-dimensional reaction system or a *surface metabolism*. So far laboratory work has provided partial support for this hypothesis.

14.2.2 Earth Formation, Origin and Early Evolution of Life

Primitive forms of life likely appeared on Earth before 3.5 billion years ago, which is the approximate age of the oldest (unanimously accepted) biosignatures: stromatolites (Fig. 14.1). The morphological variety of Archean stromatolites provides evidence that early life had developed, evolved, and diversified into different environmental niches by that time. Whether those life forms that inhabited early Earth made our planet a unique place when compared to other rocky planets and hypothetically habitable moons, asteroids, and comets is a key challenge in astrobiology.

The emergence and increase of life's complexity required a planet with specific physical and chemical conditions (habitability) that could sustain chemical systems capable of Darwinian-type evolution. During the earliest stages of evolution on Earth, primitive life-forms would have opportunistically adapted and developed suitable strategies to take advantage of the energy and nutrient sources available in a variety of environmental niches. It should be noted that the conditions in which life appeared and first evolved were *extreme* compared to those of the present-day Earth. The first billion of years of transformative planetary forces (e.g., Chap. 6)

Fig. 14.1 Approximately 3.5 Gyr old stromatolites (cross-section view), Western Australia. Source: B. Cavalazzi

must have deeply influenced the evolution of life. In spite of the fact that it is still not known how primordial life emerged from a non-living world, parts of that puzzle may be solved by reconstructing the primitive environmental conditions preserved in the terrestrial sedimentary record. By comparing the possible evidence for life in those deposits with their counterparts in modern ecosystems, astrobiologists can improve the ability to decipher the ancient paleobiological record on Earth and, potentially, on other rocky planets that became inhabited by microbial-type life. Theoretical and experimental approaches can also be employed to evaluate how geochemical processes and solar and galactic cosmic radiation alter biosignatures that might become exposed on the surface of other planets.

It seems likely that some geological evidence of early niches for life will be assessed through space exploration of other planetary bodies in our Solar System. There are regions on Mars, for example, where unaltered rocks much older than the oldest rock deposits preserved on Earth are available and exposed at the surface. This type of early rock record may harbor evidence that could elucidate some aspects of the origin of life or nature of early life that are simply missing on Earth because the earliest crust has been processed and recycled extensively by plate tectonics (Chap. 10).

14.2.2.1 Hadean Habitable Earth

Earth and the Moon formed about 4.5 Gyr ago (Fig. 14.2) as a consequence of a Mars-sized body (*Theia*, Chap. 11) colliding with proto-Earth. Immediately after its

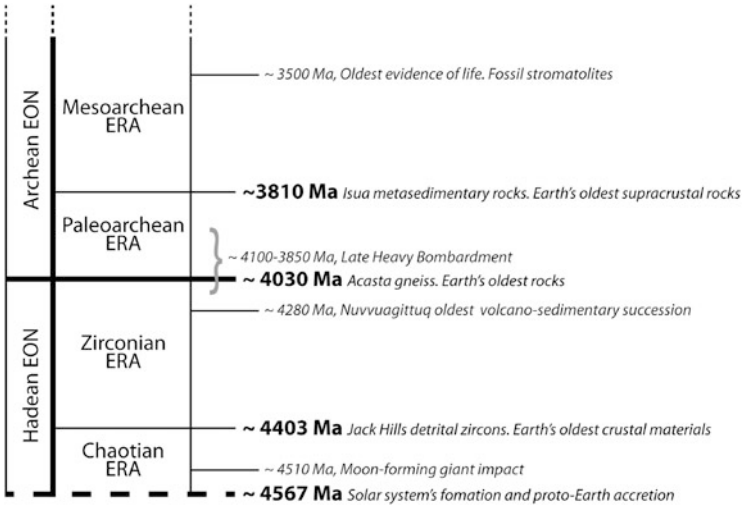

Fig. 14.2 Simplified scheme of the major steps in the formation of the Earth and its evolution to a habitable planet. Source: based on the time-scale division by van Kranendonk et al. (2012), modified from Cavalazzi and Barbieri (2016)

formation, the newly formed incandescent and chemically dynamic planet began to cool and differentiate. Radiation from the Sun should not have significantly influenced the thermal balance of the evolving planet, though the presence of a closer Moon (which caused Earth to have a faster rotation rate around its axis) should have influenced Earth's cooling history, and, ultimately the occurrence of life and its evolutionary direction. A possible super-greenhouse effect (from the high atmospheric concentrations of CO_2 and/or organic molecules produced by ultraviolet radiation reacting with the large amount of CH_4) would have occurred as a consequence of the formation of an early atmosphere and oceans. As there is no rock record from this time on Earth, the major planetary evolutionary events, such as cooling and differentiation of the crust and mantle, establishment of the Earth's hydrological cycle, and the formation mechanism and composition of its atmosphere, are difficult to constrain (Chaps. 10 and 11).

In spite of the lack of preserved crustal regions of Hadean age, 4.4 billion year old zircons from the Jack Hills region of the Pilbara Craton in Western Australia—the oldest known materials of Earth's crust—indicate that it began to differentiate during a very early stage of the Hadean, that crustal fractionation was occurring (Chap. 10), and that a possible liquid global ocean was established by that time. These ancient minerals may have formed as a result of a fast crustal recycling mechanism and fractionation of hydrated crust indicating the presence of liquid water on our planet at that time. Hadean Earth probably had a surface covered by highly saline, liquid ocean bodies shrouded by a dusty and dense, reducing atmosphere. The presence of liquid water enhanced the habitability potential of the planet and is one of the essential ingredients for life (especially when coupled with

suitable environmental niches that have sustainable carbon and energy sources, like hydrothermal systems).

Early Earth can be envisaged as an ocean-dominated, potentially habitable planet where thermophilic or hyperthermophilic life appeared and thrived in hydrothermal and/or deep, protected subsurface ecological niches very early in Earth's history (Chap. 11). Although the Hadean Earth could have been habitable relatively soon after its formation, the dramatic impacts during the debated Late Heavy Bombardment (LHB), between 4.1 and 3.85 Gyr ago it, if it occurred, could have affected our planet in numerous ways. Some hypothesise that a part, if not all, of the oceans could have evaporated, which would have seriously compromised the Earth's habitability. In other scenarios, only the upper layers of the ocean would have been evaporated. If life had already emerged before the hypothesized LHB, would it have been completely eradicated? As we know, microbial life is very difficult to eliminate once it has taken hold. Even in a worst-case scenario, microbial cells would have survived in crustal fractures or hydrothermal structures at the seafloor. Indeed, the 16S rRNA-rooted tree of life indicates that the earliest organisms were hyperthermophiles. It has been suggested that this may be an artifact of primitive cells surviving the LHB in hydrothermal systems.

14.2.2.2 Post-LHB Inhabited Earth

Whether or not the LHB took place (excursion in Chap. 7), the Early Archean Earth was a potentially habitable planet. There is evidence of life (in the form of geochemical signals of highly isotopically fractionated ^{13}C) in the ~3.8 Gyr old altered sedimentary rocks of Isua Supracrustal Belt in Greenland. However, because these rocks have experienced high-grade metamorphism and their original geochemistry has been modified, the evidence presented for early life, based on them, is controversial. The primacy of this site, however, makes the Isua Belt a place of unique significance to search for evidence of the origin of life on Earth regardless of the frustrating results obtained to date. The abundance and variety of the paleobiological records in the silicified sedimentary sequence of the Pilbara Craton (Western Australia) and Barberton greenstone belt (South Africa) that range in age from 3.5–3.0 Gyr old (Fig. 14.2), as well as the variety of their paleoenvironments (based on reconstructions), reveal that life was flourishing and well established on Earth before 3.0 Gyr ago (Fig. 14.3). There it persisted and evolved from microbes to eukaryotes to human civilization over the remaining 3.5 Gyr of Earth history.

Fossil stromatolites 3.5 billion years in age are reported from the Pilbara Craton in Western Australia (Fig. 14.1). Here, in the 3.49 billion year old Dresser Formation, hydrothermal-related stromatolites are the likely result of the activity of hyperthermophilic microbes. For the stromatolites of the 3.43 billion year old Strelley Pool Formation, a shallow marine environment under the influence of hydrothermal activity has been hypothesized. In the 3 billion year old Pongola Supergroup, South Africa, marginal marine environments evidenced by tidal

Fig. 14.3 Giant, elongate stromatolite domes (cross-section view; first author for scale) from the ∼2.5 Gyr old Lyttleton Formation, Malmani Subgroup, Transvaal Supergroup. Source: B. Cavalazzi

channels have been proposed as an environment that supported the formation of columnar stromatolites by oxygenic photoautotrophic cyanobacteria.

Well-preserved ∼3.4 billion years old filamentous microfossils and ∼3.3 billion year old microbial mats of presumable photosynthetic origin have been described from carbonaceous cherts of the Barberton greenstone belt, which belongs to the Swaziland Supergroup, South Africa. Tubular morphologies attributed to biotic alteration have also been described in ∼3.4–3.3 Gyr old pillow lavas of South Africa and Western Australia. Also from the Barberton greenstone belt, in tidal deposits of the Moodies Group, ∼3.2 Gyr old, microbially induced sedimentary structures have been described.

The oldest fossil evidence of life suggests that the early Archean was inhabited by relatively diverse forms of life. The geochemical proxies, especially the mass fractionation of sulfur isotopes by atmospheric processes, suggest that the early environmental conditions were anoxic and evolved to oxic during the so-called *Great Oxidation Event*, which occurred around 2.4–2.3 Gyr ago, or even earlier. These geochemical proxies for life also suggest that earliest microbial ecosystems were sulfur-based.

The origin of life and the progressive increase of its complexity required a permanently habitable planet able to support and sustain life. Life as we know it could never have formed in an oxygenated environment because the organic and inorganic ingredients of life would have been oxidised. The timing of the mutation

that led to the production of oxygen by cyanobacteria is still not known, though their contribution to the oxygenation of the environment at the end of Archean was a necessary prerequisite for the evolution of more complex life that included the eukaryotes and, ultimately, intelligent life. The oldest known fossil record for photosynthetic oxygen dates back to ~2.95 Gyr ago, whereas the fossil record for large animals appeared 580 Myr ago when the level of oxygen exceeded 3% by volume.

14.2.3 Life and Extreme Environments

The search for life on other planets relies on knowledge of our own planet and the environmental and ecological systems that we are able to compare—at some level—to extraterrestrial environments. As space exploration continues to expand and the missions increase in complexity, they provide vast amounts of data for us to advance our understanding of the conditions that may be considered habitable on other planets. Most of the potentially habitable environments identified on other planets are similar to extreme environments on Earth. The environments of these extreme ecosystems and the organisms that inhabit them (*extremophiles*) are frequently used as model systems to refine strategies for life detection and characterization elsewhere.

Concurrent with space exploration, the exploration of extreme environments on Earth has revealed an incredible amount of microbial diversity and expanded the borders of *what is habitable* to include almost any environment that contains *available* water (i.e., water activity 0.5–1.0). Life has been found at extremes that include pH 0.5 to >11, temperatures of −20 to 121 °C, pressures up to 400 MPa, high ionizing radiation, low levels of nutrients, low water activity, and high-salinity (as high as 5.2 M). In the search for life on other planets, it must be kept in mind that the focus on carbon-based life associated with water as an essential medium for its support and propagation reflects a natural bias based on our knowledge of Earth life.

The use of Earth-based extreme environments (ecosystems and their fossilized deposits) as analogs for life detection may be highly relevant to astrobiology, though they may not hold the answers necessary to address the origin of life on other planetary bodies. Microbes have adapted the ability to tolerate and thrive in a wide range of environmental settings that, in general, may not reflect the prebiotic system or the progression of evolution that occurred as life emerged or evolved.

14.2.3.1 Thermophiles and Hyperthermophiles

Microorganisms thriving at temperatures between 45 and 80 °C are considered *thermophiles*, while *hyperthermophiles* grow at and above 80 °C (Fig. 14.4). Thermophiles and hyperthermophiles are commonly observed at deep-sea and on-

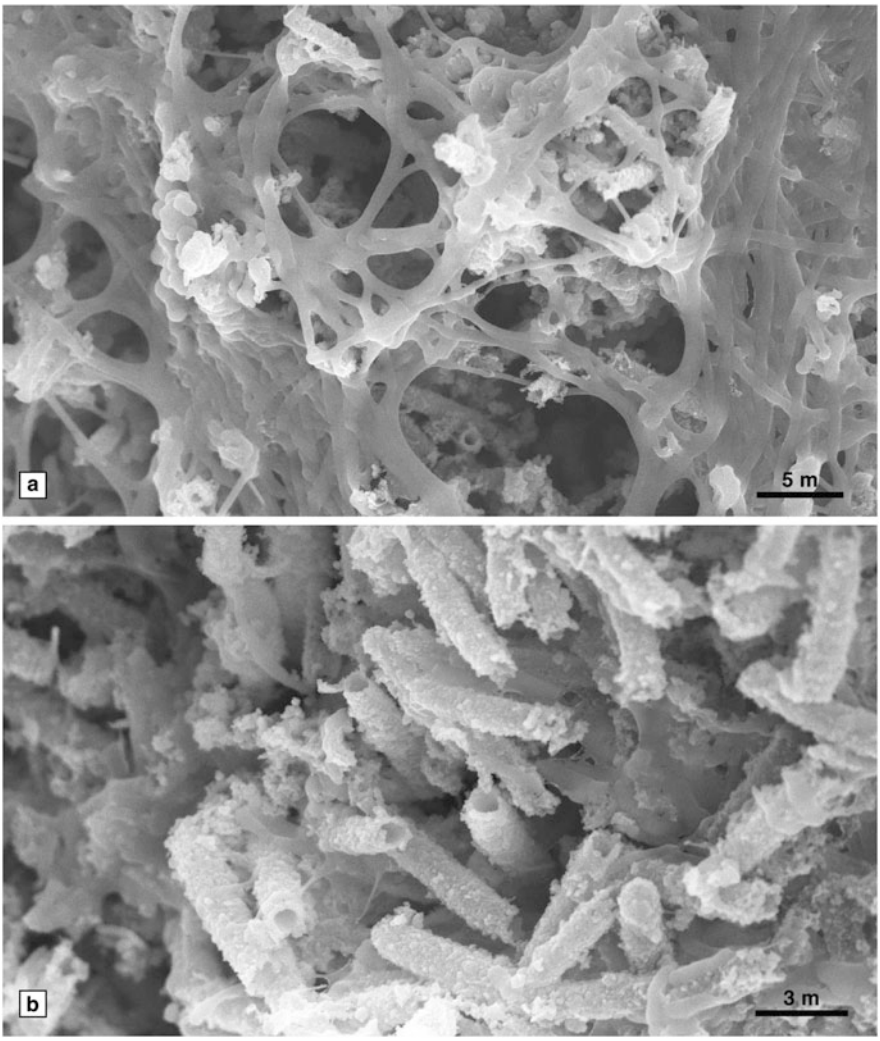

Fig. 14.4 Scanning electron microscope images of (**a**) hyperthermophilic biofilm and (**b**) cast of microbial filaments from Queen's Laundry hot spring, Yellowstone National Park, U.S. Source: B. Cavalazzi

land hydrothermal vents that include geothermally heated mud pools and soils (Fig. 14.5). As hydrothermal environments are a product of subsurface geological processes, they commonly consist of reducing fluids with low pH that are transported to their surface and near-surface environments. Consequently, many thermophilic organisms are poly-extremophiles.

Molecular biochemistry of thermophilic and hyperthermophilic adaption to life at hydrothermal vents has been studied more extensively than any other

Fig. 14.5 Dallol hydrothermal field (water temperature-pH-salinity up to respectively ∼100 °C, <1, ∼400–500 g/L), Ethiopia; the Dallol hydrothermal field is probably the most remote and inhospitable environment of our planet. Source: B. Cavalazzi

extremophilic group. At high temperatures, essential biomolecules will degrade (e.g., enzymes will denature and lose their function; membranes will become more fluid and cause cells to burst), which is why thermophiles have a number of special adaptations. Some thermophilic bacteria, for example, have more thermally stable N_4-methylcystosine rather than 5-methylcystosine in their DNA. Other thermophiles have membranes that consist of lipids with saturated and straight chain fatty acids, which allow them to maintain membrane fluidity necessary for growth and proper membrane function at high temperatures. Additionally, some thermophilic proteins have a greater number of large hydrophobic residues, disulfide bonds, and ionic interactions, while other denatured proteins may refold and repair through the action of chaperones.

14.2.3.2 Psychrophiles

Organisms that grow at or below 0 °C and those that grow optimally at 15 °C are considered *psychrophiles*. They are commonly found in a range of cold environments, from cold deep-sea to frozen on-land environments (Figs. 14.5 and 14.6). Oceans cover most of Earth's surface. Deep-sea waters usually have temperatures around 2 °C, which makes cold, saline environments common and widespread on our planet. When seawater freezes at and near the Polar Regions, the salt content of the remaining fluid concentrates and the freezing point decreases. Under such conditions, the freezing point of water has been recorded to be as low as −20 °C. Microbial metabolism has been detected at temperatures as low as −10 and at

Fig. 14.6 (**a**) Ny Alesund bay, Svalbard's glacier, Norway (source: K.O. Storvik); (**b**) Playa, Lake Lucero, White Sands National Monument, New Mexico, U.S.. Source: M. Glamoclija

−20 °C. This natural setting has led to microbial adaption at low temperatures and high-salt conditions, which makes these psychrophiles halophilic as well.

To overcome environmental constraints related to low temperatures and frequently increased salinity, which cause enzymes to become rigid and solute concentrations to increase and frequently become toxic, psychrophiles have membranes with increased levels of unsaturated fatty acids that help regulate their membrane fluidity. These organisms also produce enzymes that support transcription and translation at low temperatures. Additionally, antifreeze proteins have the ability to bind to ice crystals through their large complementary surface, which prevents ice crystals from tearing the cell membrane.

14.2.3.3 Acidophiles

Acidophiles grow optimally at low pH-values of 2. In nature, acidophiles are commonly found in hydrothermal vents and at acid mines and associated drainage fields. Interestingly, many acidophiles will cause and/or contribute to the acidity of their habitat by, for example, oxidizing sulfur and producing sulfuric acid as a by-product. To survive in such an environment, acidophiles maintain their intracellular pH at 5 or 6 by actively pumping protons out of the cell. These organisms may also have more amino acids with neutral side-groups.

14.2.3.4 Alkaliphiles

Organisms that grow optimally at pH 9 or 10 and do not grow at circumneutral pHs are alkaliphiles. The environments with high pH levels are usually serpentinization sites, soda lakes, playa deserts, and some high deserts (such as deserts in the western US, Fig. 14.6b). In alkaline environments, microbes have difficulties using ATP-synthase to produce energy and essential ions. To overcome such challenges these microbes pump their essential ions (e.g., magnesium and calcium) into the cell and export others so that they can maintain intracellular pH near neutral values. Alkaliphiles are often isolated from neutral environments with high salt concentrations (NaCl), so in addition to being alkaliphiles they are also oftentimes halophiles as well.

14.2.3.5 Halophiles

Halophilic organisms grow in fluids with high salt concentrations that can reach saturation levels; some of these organisms can even survive within salt crystals. Halophiles have been isolated from a wide range of environments that include salt marshes, subterranean salt lakes, salt quarries, and cold polar hypersaline lakes in Antarctica, the later of which is characterized by sufficiently high salt concentrations that prevent the lakes from freezing even at −20 °C. Microbes adapt to these environments by maintaining high intracellular concentrations of KCl, glycerol, glycine, mannitol, and other highly soluble compounds that maintain their osmotic balance with their environment. Though not yet understood, halophiles likely have sophisticated protein adaptations that are, potentially, similar to those of hyperthermophiles.

14.2.3.6 Piezophiles

Microbes that prefer high-pressure environments, rather than atmospheric pressure, are known as piezophiles. As with several other extremophile types, these organisms have more than one adaptation to extreme conditions. For example, piezophiles have been found at the seafloor (2 °C) at oceanic depths greater than 10 km where the hydrostatic pressure is equivalent to 110 MPa. Piezophiles have also been studied in Earth's deep subsurface where they were isolated from the environment at 40 MPa and temperatures above 100 °C. Though such environments seem harsh by anthropomorphic standards, they actually occupy a large portion of Earth's habitable environment and are relatively stable, low nutrient habitats. Piezophiles usually have very slow growth rates, which contributes to the fact that they are the least understood of the extremophiles. It is still unclear, for example, whether high-pressure adaptations require a change of regulatory modulations by one or a few genes or if it requires a global alteration of many genes in a genome. To date, the presence of monounsaturated fatty acids in membranes and of regulatory membrane

proteins OmpH/OmpL have been shown to be piezo-adaptations. Interestingly, the family of OmpH/OmpL proteins is involved in the regulation response to high temperature, *pH*, and osmolarity.

14.2.4 Astrobiology Research in Our Solar System

While Mars is the most similar to Earth, our Solar System is composed of other planetary bodies that may harbor potentially habitable environments. At 460 °C, the surface of Venus is the hottest among the terrestrial planets. Venus experiences a powerful greenhouse effect due to its proximity to the Sun and has an atmosphere composed of more than 96% CO_2. Although these attributes render the Venusian surface uninhabitable by life as we know it, the lower and middle cloud deck of the planet is hypothesized as potentially habitable because of the atmospheric temperatures (between freezing and 120 °C), the presence of liquid water, and relatively recent discoveries of microbial life at high altitudes in the terrestrial atmosphere. The planetary similarity between Mars and Earth is best illustrated by similarities in their early planetary evolution and the fact that both planets had established early hydrological cycles. There is a large number of terrestrial analogs that scientist investigate in order to assess a large diversity of Martian environments and the environments that existed at different points during the Martian history. Hence it is not surprising that Mars has been the main target of astrobiology space missions with key questions being If life emerged on Earth, did it emerge on Mars? If not, why not? Existence of life on the Martian surface today seems unlikely because of the harsh conditions (high UV radiation, lack of magnetosphere, cold and dry environment). However, during Martian early history, Mars seemed to have had a hydrological cycle that involved the presence of surface water, and it might have had a wide range of Earth-type habitats from glaciers, fluvial and lacustrine settings to hydrothermal environments (Chaps. 8 and 10). If life has had a chance to start on Mars and diversify, Mars might have had similar chances to support life as Earth did.

While the use of Earth-based analog environments has been an important factor in preparing for astrobiological exploration of Mars, other planetary bodies may harbor potential habitable environments that do not have terrestrial analogs. In those cases, experimental approaches will be needed. Icy Moons, especially Europa, Enceladus, Ganymede and Titan, provide intriguing possibilities as potential astrobiological targets beyond Mars and Venus. Active geology with evidence of exchange between surface and subsurface chemistries during the past as today too, enhances their potential to harbor potentially habitable environments. While icy moons are very intriguing, they are vastly different from any terrestrial environment and there are no adequate terrestrial analogs that could help in testing the hypothesis of potential habitability. Enceladus captured the interest of the astrobiology community when active jets of water vapor, simple organics, volatile species (N_2, CH_4, CO_2), and ice particles were detected emerging from the surface

of its south polar region. Based on composition of the ice particles and plume it has been hypothesized that Enceladus has a sub-ice alkaline ocean, which presumably has the characteristics deemed necessary to support life (e.g., such an ocean would be characterized by liquid water, bio-essential elements, and a number of possible nutrient and carbon sources). Any terrestrial analogs for sub-ice habitats on Enceladus would likely be similar to astrobiology analogs for Europa. The temperature (−187 to −141°C) of the ice shell on Europa is significantly lower than the coldest temperatures at which terrestrial life can live. Nevertheless, it could act as a shield for organic compounds and life forms, should they exist, in a sub-ice ocean from damaging surface radiation. Although the chemical composition of Europa's sub-ice ocean is not known, deep-sea hydrothermal vents and deep-sea sediments have been proposed as possible terrestrial analog sites for this icy moon. Titan is the only moon in our Solar System that possesses an atmosphere and geologically dynamic surface characterized by evidence of fluvial, lacustrine, and marine environments. Titan hosts hydrocarbon-rich lakes that contain a liquid mixture of methane, ethane, and heavier hydrocarbons (i.e., a liquid, carbon-rich environment at thermodynamic disequilibrium), and many have hypothesized subsurface oceans that presumably consist of water and ammonia. The search for life in such different settings may be best aided by experimental approach, rather than looking directly and solely in natural terrestrial settings.

14.3 Planetary Exploration

Ongoing discoveries of (i) life on Earth in nearly every type of extreme environment where liquid water and metabolically suitable carbon, energy, and nutrient sources exist; (ii) liquid water on or beneath the surfaces of planets, moons, and asteroids; and (iii) rocky, Earth-like planets in habitable zone orbits around young Sun-like stars, strengthen the case for extraterrestrial life to exist beyond Earth.

14.3.1 Space Exploration of Mars

The likelihood that fossil life will be discovered on Mars increases with each subsequent mission as more information is learned about where to look (i.e., in potentially habitable geological deposits) and what to look for (i.e., possible biosignatures). Our understanding of the habitability potential of the planet has developed via the findings of multiple missions that sequentially investigated the surface of the planet remotely and in situ over the past four decades (Chaps. 3 and 5).

Launched in 1971, Mariner 9 was the first spacecraft to orbit Mars. Equipped with a camera and spectrometers, it collected exceptional data of Mars' atmosphere and surface, which changed our perception of the Red Planet. Geologic features

(e.g., canyons and valley networks) on the surface of Mars provided the initial evidence that water once flowed on its surface. This discovery provided the momentum needed for the development of the Viking program.

The Viking Mars mission, which sent two landers that made soft landings on the surface of the planet in 1976, was an ambitious astrobiology mission in that all of the Viking experiments were designed to detect biosignatures in Martian soil (Chap. 5). Although one of the experiments tested positive for life, the results were generally dismissed because they were inconsistent with the results of the other experiments. Subsequent studies have shown that the capacity of the instruments to detect life was limited by their resolution and because so little was understood about the surface geochemistry and its effects on organics.

In 1996, the discovery of possible biosignatures in ALH84001, an ancient Martian meteorite (Chap. 6) that was collected in Antarctica in 1984, reinvigorated the search for life on the red planet. The controversial findings also illustrated the challenges inherent in searching for microscopic-to-submicroscopic morphological and chemical evidence of extraterrestrial life in rock.

In 1997, the NASA Astrobiology Institute was established, and a timeline and budget strategy was developed to launch a sequence of future missions to Mars (e.g., Sojourner, Mars Exploration Rovers (MER), Mars Science Laboratory (MSL), ExoMars and Mars2020, ExoMars 2020). With respect to the Solar System, the recent Decadal Survey of NASA 2015–2025[2] and the Cosmic Vision survey of ESA 2015–2020[3] have the same underpinning *big questions* related to the formation and habitability of the Solar System and how life emerged. It is within this framework that future missions are planned. For Europe and ESA, ExoMars 2020 falls within the long term goal of eventually having a human presence on the red planet (2040–2050?). In order to send humans to Mars, it is necessary to determine whether the planet had been inhabited (or is still inhabited in the subsurface) and be able to return materials from the surface of the planet. ESA will work with NASA on its Mars 2020 mission to bring samples back from Mars. Bringing samples back from Mars is a complicated issue and will require a series of coordinated missions (excursion in Chap. 5).

Recent Mars obiter and rover missions have confirmed that extensive quantities of liquid water existed on the surface of Mars in its past. In particular, the Phoenix lander confirmed that ice is buried below the first few centimetres of the Martian soil in the Northern plains, while the Mars Exploration Rover mission found mineralogical evidence of the past presence of liquid water in two very different geologic settings. More recently, the Mars Science Laboratory has detected organic molecules and other chemical compounds that would have made biotic evolution possible on the water-rich early Mars.

[2] http://science.nasa.gov/earth-science/decadal-surveys/.

[3] http://www.sci.esa.int/cosmic-vision/.

14.3.2 Biosignatures and Life Detection

Among the central concerns in astrobiology is life detection. As the earliest and dominant life forms for nearly 7/8 of Earth's history were microbial communities, knowledge of their life styles and fossilization potential in extreme environments considered Mars analogs guide life detection strategies for astrobiology missions. Evidence of past or present life that could not have formed unless life was present is commonly referred to as *biosignatures*. Examples include (i) primary (e.g., pigments) and diagenetically altered (i.e., lipids) biosynthetic molecules; (ii) carbonaceous cellular and extracellular remains that retain biologically fractionated isotope patterns; (iii) biologically concentrated elements, biominerals, and mineral assemblages in chemical disequilibrium due to biological activity; (iv) non-racemic concentrations of chiral molecules; and (v) microbially influenced sedimentary fabrics and structures (e.g., stromatolites).

The search for ancient life on Earth has demonstrated the importance of locating fossilized signatures of life evidenced by multiple biosignatures; detection of such evidence in more than one sample; confirmation of any discovery by more than one research group and, in some cases, proving experimentally that the evidence for life is exclusively of biological origin. Extensive planning for astrobiology mission instrument payloads by multiple space agencies includes the design and development of miniaturized instruments capable of detecting and characterizing potential biosignatures. A more recent development in astrobiology search strategies include the development of systems capable of cashing samples for planned sample return missions, where more sophisticated instrumentation in laboratories on Earth can be used to interrogate potential signatures of life.

14.3.3 Planetary Protection

Of critical importance since the beginning of the space age and exploration of other worlds is the need to investigate them in such a way as to avoid disruption and contamination of possible zones of life. Planetary protection is an overarching goal during mission design for all of the major space agencies that send spacecraft to other planetary bodies so that they can be explored responsibly during conditions of extreme uncertainty.

This concern was demonstrated early in the history of space exploration: the members of the Space Studies Board, established in 1958, defined the first protocol for Planetary Protection (e.g., sterilization of the spacecraft before launch and quarantine for samples returned from other planets) that NASA still maintains today. COSPAR and NASA recommendations for planetary protection protocols in the Apollo program focused on prevention of back contamination and exposure of Earth to organisms or chemicals from the Moon. Planetary protection protocols for the Viking Mars lander missions focused on preventing forward contamination of Mars

because it involved a rigorous scientific investigation that included a critical search for life. The combination of planetary protection issues that surfaced during the design of these missions became a guiding principle that improved the design of all subsequent and future interplanetary missions to other (potentially habitable) worlds.

Current planetary protection policy addresses procedures and technologies for contamination control that depend upon the type of space mission. NASA's Office of Planetary Protection has defined five planetary protection levels. Each mission is therefore categorized based on the type of encounter it will have (flyby, orbiter, rover) and the nature of the target (planet, moon, asteroid, comet). If the target body has the potential for life detection or evidence for prebiotic chemistry, the spacecraft and its instrument payload will need to satisfy a higher level of cleanliness and operating restrictions will be mandatory.[4]

Planetary protection protocols are critical priorities for near term planned and future possible space missions, especially given the possibility that multiple samples could be returned to Earth as part of a future Mars Sample Return mission and a human mission to Mars is envisioned. Updated protocols for forward protection, life detection in extraterrestrial samples, and proper handling of samples in terrestrial laboratories are all of great importance. The best sample to be returned to Earth is an uncontaminated one, which means rigorous protocols must be in place to eliminate the possibility of bringing extraterrestrial life and chemistry to our planet (and if we do, we need to be able to discern them). A sample return mission must also include rigorous protocols for sample handling and storage that will maintain sample integrity and cleanliness and, thereby, protect the Earth.

14.3.4 Life in the Universe

When searching for evidence of life beyond our Solar System, we seek first to understand the potential for habitability of the celestial body, that is the potential to develop and sustain life, as defined largely through extrapolation of known terrestrial conditions and their combination (e.g., geophysical, geochemical, and astrophysical factors) that allowed for the emergence of life on Earth. Since liquid water is required for life as we know it, a habitable planet will likely be Earth-like, that is, with liquid water at its surface. It is worth noting that all planets orbit a central star(s) in a defined star-lifetime, yet not every planet in the habitable zone of a central star is habitable. Therefore, in the search for life beyond our Solar System, astrobiologists focus on the details of investigations of different planetary environments that could support liquid water and then assess the potential for past or present habitability.

[4] http://planetaryprotection.nasa.gov/categories.

Stars are a key factor in defining the Earth-like habitability potential for a planetary system. Stars with lifetimes of at least a few billion years may provide life on a particular planet with the time and opportunity to evolve, and only those stars whose spectrum falls mostly into the visible range can allow processes such as photosynthesis. This spectral range probably accounts for 5 to 10% of stars in the local Milky Way Galaxy. Whether fainter red dwarf stars are also suitable hosts for habitable planets is an important, yet open, question, as they are prevalent in the stellar population. The circumstellar habitable zone is the region around a star within which a planetary object, with sufficient atmospheric pressure, can support liquid water at its surfaces. The limits of the habitable zone are calculated by using the known parameters of Earth and its position in the Solar System.

All stars increase in luminosity as they age, and as a consequence their potential habitable zone will migrate outwards. If this process happens too quickly (e.g., in case of a super-massive star), planets may only have a brief time to reside within the habitable zone and a smaller chance to develop life. The existence of a habitable zone also requires that no large-mass body, such as a gas giant, is present within it or relatively near it in that this would likely prevent the formation of an Earth-like body. The matter in the asteroid belt, for example, appears to have been unable to accrete into a planet due to orbital resonances with Jupiter.

In the Solar System, the inner planets are terrestrial, and the outer ones are gas giants, but discoveries of extrasolar planets suggest that this arrangement may not be at all common. Nor is the model of a single central star in a Solar System necessarily the norm. The chief assumption about habitable planets is that they are terrestrial. Such planets, roughly within one order of magnitude of Earth'mass, are primarily composed of silicate rocks and have not accreted the gaseous outer layers of hydrogen and helium found on gas giants. That life could evolve in the clouds surrounding giant planets has not been decisively ruled out, although it is considered unlikely. The natural satellites of giant planets, meanwhile, remain valid candidates for hosting life. A number of candidates for Earth-like planets have been reported in recent years. Unfortunately, some of the stars were determined not to exist anymore (e.g., Gliese 581g and 581d). The confirmation of other Earth-like candidates is still ongoing.

Take-Home Messages

Astrobiology studies the origin, evolution and distribution of life in the Solar System.

Hydrothermal systems are investigated as suitable setting for early synthesis of prebiotic organic compounds

Environments characterized by extreme physical and chemical conditions on Earth can provide analogue models for potential extraterrestrial habitable environments, useful for searching and detecting life elsewhere. Different organisms adapt to specific extreme environments.

Solar System objects provide a wide range of environments with the potential of having hosted past life or harboring extant ones, including the likely presence of large subsurface liquid water reservoirs in several Icy Satellites.

Detecting life, past or extant, requires the detection of chemical, morphologic, organic and sedimentary biosignatures.

Planetary exploration, particularly if aimed at searching for signs of life, has to comply with Planetary Protection rules in order to avoid contamination and disruption of planetary and terrestrial environments.

Acknowledgements This work is a contribution to the FP7-PEOPLE-2013-CIG/INACMa (BC), and to the NASA ASTEP NNX14AT28G (MG). SLC acknowledges financial support from the NASA Astrobiology Institute and from EMSL, a United States DoE Office of Science User Facility sponsored by the Office of Biological and Environmental Research, located at Pacific Northwest National Laboratory, Richland, WA. This is a contribution to COST Action TD 1308.

Suggested Readings

Cockell, C.: Astrobiology: Understanding Life in the Universe, p. 472. Wiley, Hoboken (2015)

Gargaud, M., Amils, R., Quintanilla, J.C., Cleaves, H., Irvine, W., Pinti, D., Viso, M.: Encyclopedia of Astrobiology, p. 1853. Springer, Berlin (2011)

Gargaud, M., Martin, H., López-García, P., Montmerle, T., Pascal, R.: Young Sun, Early Earth and the Origins of Life. Lessons for Astrobiology, p. 301. Springer, Heidelberg (2012)

Kasting, J.F.: How to Find a Habitable Planet, p. 326. Princeton University Press, Princeton (2010)

Knoll, A.H.: Life on a Young Planet: The First Tree Billion Years of Evolution of Earth, p. 288. Princeton University Press, Princeton (2003)

Knoll, A., Canfield, D.E., Konhauser, K.O.: Fundamentals of Geobiology, p. 456. Wiley, Hoboken (2012)

Konhauser, K.O.: Introduction to Geomicrobiology, p. 440. Wiley, Hoboken (2006)

Longstaff, A.: Astrobiology: An Introduction, p. 466. Taylor and Francis, Abingdon (2014)

Meltzer, M.: When Biospheres Collide: A History of NASA's Planetary Protection Programs, p. 522. National Aeronautics and Space Administration, Washington (2010)

Ward, P., Brownlee, D.: Rare Earth: Why Complex Life is Uncommon in the Universe, p. 338. Springer, New York (2000)

Chapter 15
Space and Planetary Resources

Angel Abbud-Madrid

15.1 Introduction

As described in previous chapters, a multitude of remote-sensing missions to the Moon, asteroids, and Mars, as well as the study of samples brought to Earth by robotic and human missions (Chap. 5) and of meteorites collected on Earth (Chap. 6), have allowed us to obtain a preliminary understanding of the geological evolution of the planets and of a small number of moons, asteroids, and comets of the Solar System (Chaps. 11–13). At the same time, we have been able to obtain a glimpse of the composition and material found on these celestial bodies. This scientific knowledge has in turn allowed us to start thinking on ways to utilize extraterrestrial resources for practical applications. A variety of potential resources have been identified, including water, volatile elements, metals and minerals, atmospheric gases, solar energy, and the ultra-high vacuum and low-gravity of space.

As a first step, these resources could be transformed into useful materials at the site where they are found to extend mission duration and to eliminate the dependence from material sent from Earth. This *living-off-the-land*-approach is commonly referred to as *In Situ Resource Utilization (ISRU)*. Once local needs are satisfied and a resource recovery infrastructure is in place, a space transportation network could be developed. In the long term, if found to be economically feasible, extraterrestrial material could also be brought to Earth. These are the same steps that have been taken in the case of terrestrial resources by humans during the exploration and settling of our planet for more than 100,000 years; space is only the next destination in the quest for further exploration and expansion of our species.

A. Abbud-Madrid (✉)
Colorado School of Mines, Golden, CO 80401, USA
e-mail: aabbudma@mines.edu

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8_15

However, since the beginning of the space age, all propellants, materials, and human consumables have been sent from Earth. Unfortunately, our planet sits in a deep gravity well and considerable rocket energy is necessary to lift mass and put it into space. It takes ten times more energy to get into geosynchronous Earth orbit from the surface of the Earth than from the surface of the Moon and an estimated 10% of all near-Earth asteroids discovered to date are more accessible than the Moon, in terms of rocket fuel needed. As a result, if we are to eliminate our costly dependence from material sent from Earth, extend the reach and payload-carrying capacity of robotic missions, and allow humans to settle beyond our planet, resources for propellants, life support, materials for habitat construction, and all related infrastructure must come from space.

Making propellants from local resources can significantly reduce mission mass and cost, enabling extended human stays and fueling transportation systems for use within and beyond the planetary surface. Use of finely grained soils and rocks can serve for habitat construction, radiation protection, solar cell fabrication, and food growth. The same material could also be used to develop repair and replacement capabilities using advanced manufacturing technologies.

In order to identify the type of available resources in space, the feasibility of their extraction, and the potential benefits of their use, similar steps should be taken as when prospecting, mining and utilizing terrestrial resources.

15.2 Resource Prospecting

The first step in the process of utilizing space resources is to assess the location, distribution, composition, and quantity of extraterrestrial material available. A combination of remote sensing techniques and thorough surface prospecting using a variety of geotechnical measurements, drilling, and sampling are required to reliably detect and characterize concentrations of useful material in any planetary body.

Although valuable resources have been identified on the outer planets (helium-3 in Jupiter, Saturn, Neptune, and Uranus) and their moons (hydrocarbons on Titan and water on Europa and Triton), this chapter will concentrate mainly on the resources of the Moon, Near Earth Asteroids (NEAs), comets, Mars, and its two moons, Phobos and Deimos. From the technical and economical feasibility standpoint, these bodies promise to be the most accessible and realistic destinations in the foreseeable future.

15.2.1 The Moon

The Moon holds not only the scientific clues to understand the history of the formation and evolution of our planet, but also promises to be the first celestial body from where resources could be extracted and utilized to start losing our dependence

from terrestrial materials. As a result of its proximity to Earth, much more is known of the geology and elemental composition of the Moon than of any other body beyond our planet. Consequently, the Moon is most likely the first destination where an infrastructure based on space resources could be developed.

More than fifty successful missions have been sent to explore our closest neighbor, samples have been obtained from nine sites located on the near-side equatorial regions, and global mineralogical and geochemical maps of the lunar surface have been obtained by orbital remote sensing. Much needs to be done to get a complete picture of the nature, amount, and distribution of potential resources on the Moon, learn about the composition and physical state of its surface materials, and to understand the processes that produced and concentrated these resources. Nevertheless, significant exploration has been accomplished to put together a preliminary inventory of lunar resources.

As explained in Chap. 11, there is much evidence to support the hypothesis that the Moon was formed from the condensation of vaporized mantle material produced by the giant impact of a Mars-size object with the early Earth and that soon after a magma ocean existed on the Moon starting around 4.5 billion years ago. As the magma cooled and began crystallizing, the first minerals to form were the silicates olivine [$(MgFe)_2SiO_4$] and pyroxene [$(CaFeMg)Si_2O_6$]. As a result of their high density, these minerals sank to the bottom of the mantle. The less dense, mostly anorthositic plagioclase feldspar ($CaAl_2Si_2O_8$ with some $NaAlSi_3O_8$ in solid solution) crystallized and floated, forming a crust about 50 km in thickness. The magma ocean crystallized within 100 million years, though some remaining magmas with incompatible and heat-producing elements possibly remained partially molten for several hundred million years. These remaining magmas—called KREEP for being rich in potassium (K), Rare Earth Elements (REE), and Phosphorous (P)—finally concentrated in the Oceanus Procellarum and Mare Imbrium regions.

This early differentiation resulted on the formation of a lunar core, mantle, and crust. These processes also gave rise to the two main geological features of the lunar surface. The light-colored highlands or terrae are the original primordial crust of the Moon composed primarily of anorthositic rocks containing more than 90% plagioclase feldspar, with anorthite being the dominant mineralogy. The highlands cover nearly all of the far side and about 66% of the near side for a total of 83% of the lunar surface. The dark lowlands or maria were formed by large lava flows which covered the giant impact basins mostly on the near side. They constitute 17% of the lunar surface and are the most obvious and smoothest of all lunar features. These lowlands are composed mainly of basaltic volcanic rocks which, having originated by partial melting of the mantle, their mineralogy is dominated by pyroxene and olivine silicates, as well as ilmenite ($FeTiO_3$) and spinel ($MgAl_2O_4$) oxides, with various proportions in different lava flows. Lunar ilmenite is characterized by a Ti content higher than on Earth, with up to 8% TiO_2 in some regions.

Thus, with a few exceptions in concentration amounts, the elements that make up most of the Earth (O, Si, Al, Fe, Mg, and Ca), also make up most of the Moon. As shown on Fig. 15.1, from an elemental point of view (by mass), with differences

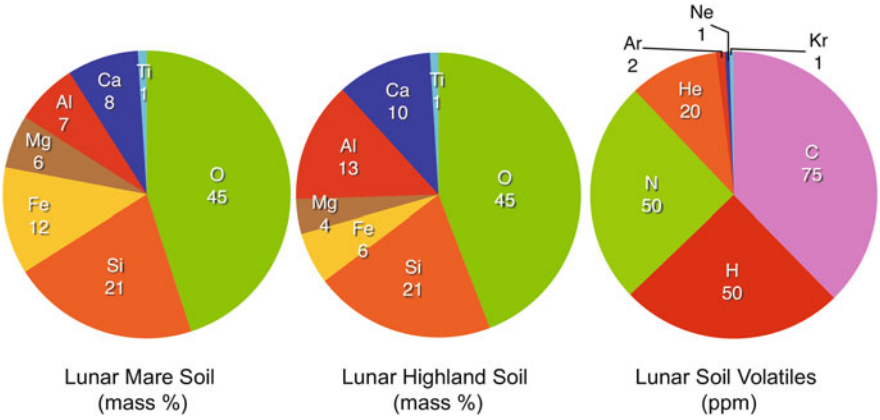

Fig. 15.1 Average chemical composition of the lunar mare and highland soils with overall concentration of volatiles. Source: data from Heiken et al. (1991)

between the highland and mare regions, the lunar surface is rich in O (45%) and Si (21%) bound in silicates and metal oxides, as well as in Al (7–13%), Fe (6–12%), Mg (4–6%), Ca (8–10%) and Ti (up to 1%). Soils derived from highland rocks have lower abundances of Fe and Mg, but tend to be rich in Al and Ca.

Other less abundant elements could also be extracted, such as Na, K, Sr and Ba from plagioclase and potassium feldspar, and Mn and Cr from pyroxenes. Chromium might also be recovered from spinel, found in mare basalts. Finally, KREEP material and phosphate minerals may be a useful source of incompatible trace elements, such as: alkali and alkali-earth elements, the Rare-Earths, B, Zr, Nb, Sn, Hf, Ta, and radioactive elements such as U and Th.

In addition to the original geochemical processes that formed the Moon's crust, as a result of billions of years of meteorite and micro-meteorite impacts on both the highlands and mare regions, the entire lunar surface is now covered with regolith, a layer of fragmental and unconsolidated rock material. The regolith is approximately 4–5 m thick in mare regions, and 10–15 m thick in the highland areas. While consisting of mostly powder-like grains in the upper surface, the regolith acquires a higher than 90% density by a depth of 30 cm. The regolith has an average grain size of about 60–80 μm and it is also composed of sub-cm rock fragments referred to as lunar soil. Individual lunar soil particles are mostly agglutinates, which are aggregates of small mineral grains, gases, and older agglutinates bonded together by glass and formed by the melting and mixing produced by micrometeroid bombardment of the lunar regolith. In general, if mineral beneficiation is to be used, rocks rather than regolith may be the best source, given that mineral fragments in regolith are less abundant than in the source rocks. Nevertheless, because the surficial, unconsolidated, and fine grained nature of regolith, it is likely that it will become the basic feedstock for most processes used for extracting and processing lunar raw materials.

The mineralogical and geochemical composition of regolith at any given location is mostly a result of the composition of the rocks beneath it. However, the regolith also exhibits concentrations of valuable elements due to external processes different from the ones responsible for the main elements mentioned above and also different from the hydrothermal and biological processes responsible for ore formation on Earth. Solar-wind implanted volatiles, water ice retention in permanently shadowed craters, enriched iron and siderophile elements on meteoritic material, and volcanic glasses with enriched volatiles on pyroclastic deposits are some of these processes known to be responsible for concentrating valuable resources.

Solar-wind particles (plasma of ionized atoms) implanted in the grains of the lunar regolith for billions of years as a result of the Moon's lack of atmosphere or magnetic field may constitute unique lunar resources. The volatile elements H, He, C, and N, as well as noble gases Ne, Ar, and in much smaller quantities Kr and Xe are ubiquitous in the regolith. As shown in Fig. 15.1, these elements have low abundances in ppm and the highest concentrations are found on the finest-grained fractions (less than 20 μm) of the regolith. The solar-wind volatiles, most notably H and He, are retained preferentially by ilmenite. In addition, as a result of the continued overturning of the regolith by meteorite impacts, solar-wind implanted volatiles are expected to be present throughout the regolith layer. As found on Apollo drill cores, their concentrations are approximately constant within the uppermost 2–3 m. It is interesting to note that while only in very low concentrations of a few ppb in the lunar soil, the isotope of helium, ^{3}He, is found in much greater quantities than on Earth's atmosphere. This is significant, given the possibility of using ^{3}He as fuel on future fusion reactors.

The Moon has been generally considered to be a dry body, as confirmed from the analysis of lunar samples. However, it was recognized as early as 1961 that water derived from the impacts of comets and meteorites with the lunar surface may have deposited as ice on cold (temperatures below 40 K), permanently shadowed regions of craters near the lunar poles. Water and OH molecules produced by the interaction between the solar wind and the regolith at lower latitudes may also migrate and become cold-trapped in the poles. As shown on Fig. 15.2, indirect evidence of ice in the form of elevated hydrogen concentrations in permanently shadowed polar craters was provided by the neutron spectrometer onboard the Lunar Prospector spacecraft in 1999 and more recently by the Lunar Reconnaissance Orbiter (LRO). Confirmation of the existence of polar ice was provided by NASA's Lunar Crater Observation and Sensing Satellite (LCROSS) mission, which impacted into a permanently shadowed region of the southern polar crater Cabeus in 2009 and measured an average water ice concentration of 5.6% (by weight) at the impact site. Observations from space and Earth-based radar instruments also indicate the presence of substantial deposits of relatively clean ice in several permanently shadowed regions.

Besides the ice found inside permanently shadowed craters, infrared measurements from orbit have detected surficial hydrated minerals at high latitudes. Whereas in the form of water or OH, these molecules (in the hundreds-of-ppm range) are produced by the reduction of iron oxides in the regolith by solar-

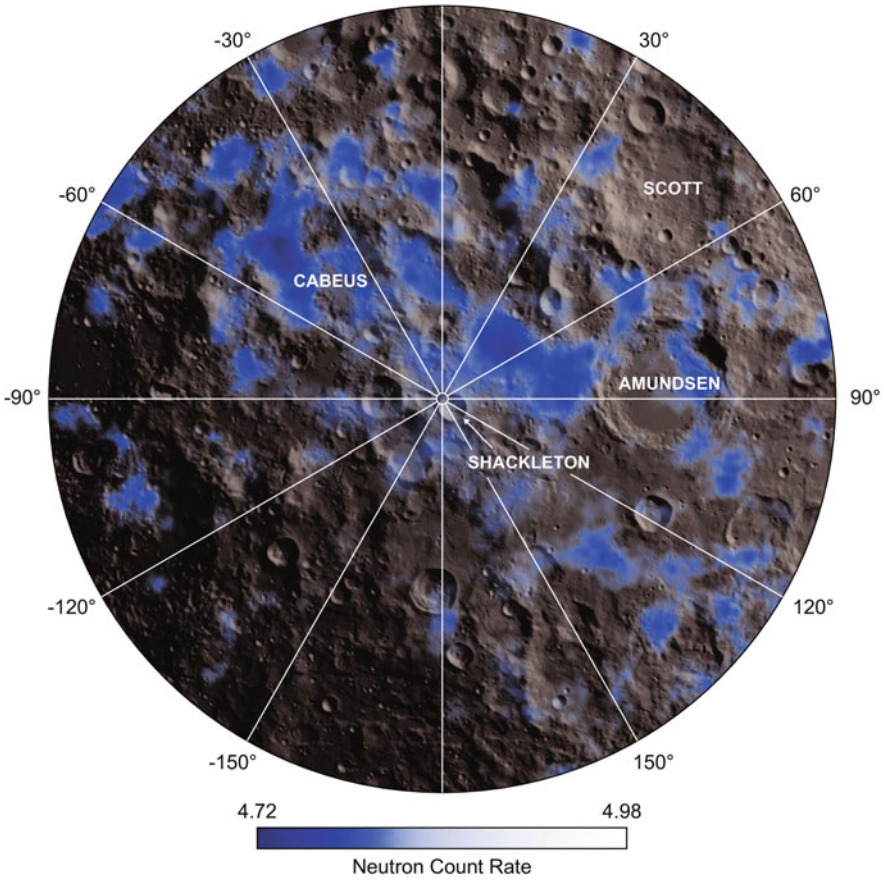

Fig. 15.2 Map of the Moon's south pole indicating the presence of hydrogen from measurements taken by the Lunar Reconnaissance Orbiter (LRO) Lunar Exploration Neutron Detector (LEND) inside and outside permanently shadow craters over two and a half years. Source: NASA/GSFC/SVS/ROSKOSMOS, Mitrofanov et al. (2010)

wind hydrogen, and retained at these high latitudes at temperatures below 110 K. In addition, hydrated minerals in the range of 5–30 ppm were found in lunar pyroclastic deposits in the Apollo 15 and 17 missions, while H_2O concentrations up to 1500 ppm were measured in similar deposits by recent infrared remote-sensing observations. Although H_2O concentrations in these regions are one to two orders of magnitude lower than the amounts detected inside permanently shadowed craters, these minerals are located in more accessible regions with direct exposure to solar energy and communications from Earth. Thus, they may represent a more economical and technically feasible option for extracting water. Finally, it is important to note that if water has been deposited in the cold traps, other volatile elements may be concentrated there as well. Comets contain abundant and valuable

volatiles (from an ISRU perspective), such as C and N, which may also be present as condensed compounds in those regions.

Although Iron (Fe) is abundant in the silicate minerals and the oxide mineral ilmenite, the most likely practical source of metallic Fe and siderophile elements, including Co, Ni, Ge, Mo, Au and the Platinum Group metals or PGMs (Ru, Rh, Pd, Os, Ir, and Pt) may be from meteoritic material. The lunar surface has been bombarded by meteoritic material by billions of years resulting in lunar soils having concentrations up to 2% of chondritic meteoritic material. Much higher localized concentrations of native Fe (Fe_0) and associated siderophile elements, could be found in the vicinity of any Fe-rich meteorites on the lunar surface. In addition, Fe_0 in the regolith is also found as vapor-deposited nanophase Fe (generally, 1 μm) within impact glass particles or agglutinates by the reduction of iron oxides by solar-wind hydrogen.

15.2.2 Mars

After the Moon, Mars is the most studied planetary object to date with close to thirty successful spacecraft missions including flybys, orbiters, landers, and surface rovers. In terms of resources, Mars has a wide variety and a large quantity of useful elements. Contrary to the Moon and asteroids, Mars has a thin atmosphere (with an average pressure of 600 Pa) composed mostly of CO_2 (approximately 96%) with other minor elements (1.9% Ar, 1.9% N_2 and traces of free oxygen, CO, H_2O, and CH_4). This is an important source of carbon and oxygen. However, the most important resource on the Red Planet for propellant production, materials processing, human consumption, plant growth, and radiation shielding is undoubtedly water.

The recent finding of carbonates on the surface implies the past existence of water on Mars. The planet presently lacks oceans, but appears to have significant water buried below the surface. While water is present in its liquid form at depths of 200–300 m and in small amounts as vapor in the atmosphere, it is abundant almost everywhere else as ice. Water ice is visible at the surface in the north polar ice cap, it is abundant beneath the permanent carbon dioxide ice cap at the south pole and is also present in the shallow subsurface (1–10 m deep) at mid-latitudes. As seen in Fig. 15.3a, the Phoenix lander in 2008 exposed a layer of ice hidden under a few centimeters of regolith at 68°N; the ice disappeared days later by sublimation. In total, more than five million cubic kilometers of ice have been estimated to exist at or near the surface of modern Mars, enough to cover the whole planet to a depth of 35 m.

Water is also locked into the structure of abundant water-rich materials, including clay minerals (phyllosilicates) and sulfates. The surface of Mars is mainly covered with basalt (Chap. 8), made up mostly of the silicate minerals olivine, pyroxene, and plagioclase feldspar. This was recently confirmed using X-ray diffraction analysis by NASA's Curiosity rover. When exposed to water and atmospheric gases, these minerals chemically turn into new minerals, some of which may incorporate water

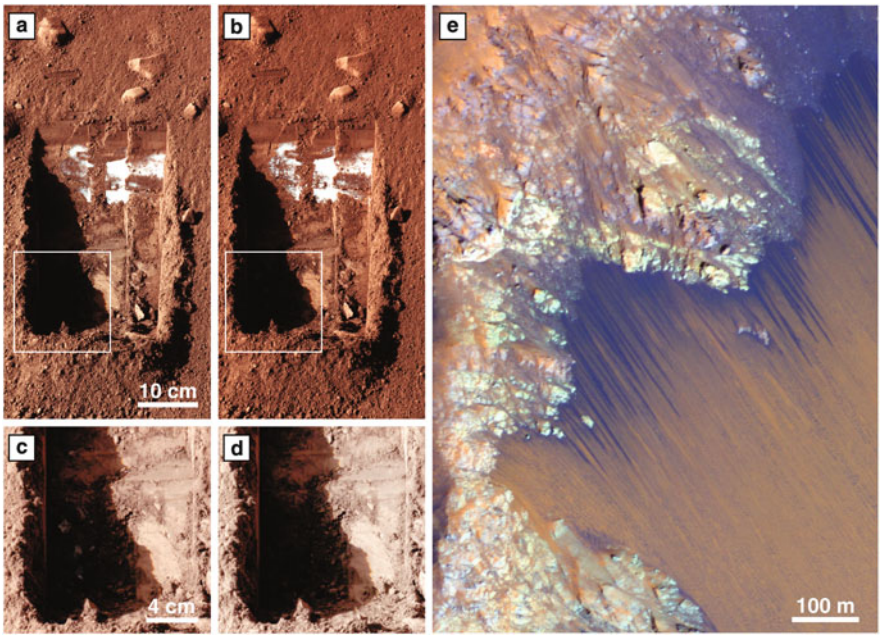

Fig. 15.3 Image of trench excavated by NASA's Phoenix Mars Lander taken on the (**a**) 21st and (**b**) 25th days of the mission, or Sols 20 and 24 (June 15 and 19, 2008) showing sublimation of subsurface water ice; see enlarged views (**c**) and (**d**) of respective lower areas. (**e**) Image from the High Resolution Imaging Science Experiment (HiRISE) camera on NASA's Mars Reconnaissance Orbiter showing the appearance of recurring slope lineae on a Martian crater wall as a result of salty water flows during the summer season. Source: (**a**–**d**) NASA/JPL-Caltech/University of Arizona/Texas A&M University, Arvidson et al. (2009). (**e**) NASA HiRISE, JPL-Caltech, University of Arizona

into their crystalline structures, either as H_2O or as the hydroxyl radical (OH). These hydrated minerals include the iron hydroxide goethite, the evaporate minerals gypsum and kieserite, opalline silica, and phyllosilicates (clay minerals), such as smectite, kaolinite and montmorillonite. Recently, the Opportunity rover confirmed the presence of Ca, S, and water in a vein of gypsum sticking out of the soil. Besides ice and hydrated minerals, recent observations by NASA's Mars Reconnaissance Orbiter of dark narrow streaks (recurring slope lineae, RSL) on the walls of Mars craters appear to indicate the presence of hydrated salts caused by seasonal flow of salty liquid water (see Fig. 15.3e). Finally, a valuable companion of water on Mars is deuterium, a stable isotope of hydrogen, which is five times more abundant on Mars than on Earth. Just like helium-3 (^{3}He) on the Moon, deuterium has also been identified as an important fuel for future fusion reactors.

Although little is known about mineralization and ore formation on Mars, exploration by surface rovers and orbital remote sensing, as well as meteorites from Mars found on Earth have revealed that local concentration of potentially high-grade

mineral ore may have been caused by hydrothermal fluids resulting from volcanic activity, large igneous province formations, and impact craters. Thus, the best places on Mars to look for valuable ores may be on volcanoes, lava flows, and impact craters. In addition, these large craters expose rocks to prospecting, conveniently saving deep excavation operations. As on Earth, elements that may be extracted from large igneous province formations include Ni, Cu, Ti, Fe, Pt, Pd, and Cr. Impact-related ores may yield Cu, Zn, Pb, Au, Ag, and Ba. Finally, mantle uprisings are also responsible of concentrating extremely valuable metals such as: Pt, Pd, Os, Ir, Rh, and Ru, and there are vast quantities of Fe, Sn, Cr, Ti, Au, and V. Mars has a very large core and very thin mantle, so mantle uprisings could be common on the planet. Finally, surface enrichment may have also been caused by mineral sand deposits as a result of strong wind activity.

The Mars landers (Chap. 5) Viking I, Viking II, Pathfinder, and the Opportunity, Spirit, and Curiosity rovers found the Martian soil to be generally silicate in composition with a high iron-mineral content along with other magnetic material, as well as high concentrations of sulfate and chloride salts and traces of bromides. As for metals, the rovers identified Al, Fe, Mg, and Ti. The Opportunity rover found small hematite-rich structures, named *blueberries*. In addition, both Spirit and Opportunity rovers found nickel-iron meteorites on the surface of Mars.

15.2.3 Asteroids, Comets, and the Moons of Mars

As described in Chap. 13, asteroids are shattered rocky and metallic bodies left over from the formation of the Solar System that never grew large enough to become planets. Some asteroids pass close to the Sun, and others travel beyond the orbit of Neptune. A large number of them are located between Jupiter and Mars, in the so-called Main Belt, as a result of the strong gravity in Jupiter. Because of gravitational perturbations and collisions, many asteroids have left the Main Belt, have crossed Earth's orbit, and even impacted our planet. In the past 20 years, more than 12,000 of these Near Earth Asteroids (or NEAs) have been found as part of a worldwide effort to inventory such objects to protect the Earth against potentially disruptive impacts. Some of these NEAs are in fact burned-out comets, which have stopped releasing volatiles.

Given their relative close proximity to Earth, their extremely low gravity force, and their wealth of resources, NEAs are considered potential targets for volatiles and metal extraction, sometimes more attractive than the Moon and Mars. Since NEAs are the most probable source of meteorites, much has been learned over the past few decades from the more than 50,000 meteorite samples that have been analyzed in laboratories. To provide a more complete picture of these bodies, telescopes are used to measure their spin rate and size, radar observations provide their three-dimensional shape, and infrared measurements provide their albedo. There have even been more than ten space missions to asteroids to date (Chap. 3). From the data

gathered from all these sources, it is believed that enormous quantities of resources can be found in these bodies.

Surface composition and potential resources of NEAs can be inferred from laboratory studies of meteorites (Chap. 6) and from the spectral reflectivity studies of asteroids at ultraviolet, visible, and near-infrared wavelengths (Chap. 3). Given that all NEAs discovered to date are below a few kilometers in diameter, these can be considered undifferentiated objects where their surface and interior composition are practically the same. Thus, based on the knowledge gathered until now, asteroid types can be classified in three main categories:

1. *C-type asteroids (carbonaceous)*: These are the most common objects with around 75% of the known asteroids fitting this category. They are dark in color as a result of the very high contents of opaque, carbonaceous material. They closely follow the elemental composition of the sun, making them some of the most primitive objects, formed at the dawn of the solar system. Our knowledge of these bodies is based on the chemical analysis of meteorites believed to come from these parent bodies, known as carbonaceous chondrites, which probably came from bodies too small to undergo substantial gravitational differentiation. Carbonaceous chondrites with large amount of volatiles and valuable resources average about 10% water in a clay mineral matrix and as water of hydration (often in magnesium salts, 5–15%), 2–5% carbon in the form of graphite, hydrocarbons and organic compounds, several percent sulfur in elemental, iron sulfide and water soluble sulfate forms, some nitrogen and other volatiles, and 5–15% magnetite.
2. *S-type asteroids (stony)*: These bodies amount to approximately 17% of the known asteroids and are a mixture of rock and metal mixed together. They are made of anhydrous rocky material, are the most abundant type in the inner portion of the Main Belt, and are often found as NEAs. It is believed they are the source of the chondrite meteorites, which are the most commonly found meteorites. They are composed of material that was heated to melting, but never separated from the rock. They consist of silicates (with large amounts of olivine, pyroxene, and plagioclase feldspar), sulphides (with the ferrous sulphide mineral troilite), and metals (such as nickel, iron, and magnesium).
3. *M-type asteroids (metallic)*: As a result of their high radar reflectivity, the rest of the asteroids (8% of total known to date) are composed of primarily metal. They appear to be the remnants of large (100 km and above) asteroids that fully differentiated into a core and mantle, underwent massive collisions, and left only the metallic cores. They are extremely dense, unlike any metallic ore bodies found on Earth. Besides large quantities of pure Fe, Ni, and Co they also have high concentrations of PGMs (Ru, Rh, Pd, Os, Ir, and Pt) and Rare Earth Elements distributed evenly throughout the asteroid bodies.

As for NEAs, about half of the kilometer-size NEA population and therefore about half of the mass of the NEAs, is believed to be carbonaceous, rich in water and carbon. Since the other half is dominated by S-type asteroids with a few percent of M-type bodies, then the non-carbonaceous asteroids contain around 20% metallic

iron-nickel-alloys, 6% of the ferrous sulphide mineral troilite with large amounts of olivine, pyroxene, and plagioclase feldspar, trace amounts of rare and valuable metals (especially PGMs), and non-metals (such as phosphorous, carbon, sulphur, arsenic, selenium, and germanium).

In addition to NEAs, certain comets could also be good candidates for volatile resources. As shown in Fig. 15.4a, small (less than 10 km), short-period comets produce a characteristic tail of gas and dust when they are near the Sun. Thus, they may be composed primarily of volatiles such as H_2O, CO_2, CO, NH_3, CH_4 and a wide variety of organic compounds that could be captured by an orbiting spacecraft.

Finally, although not strictly classified as asteroids, the moons of Mars, Phobos and Deimos have been found to be very much in common with carbonaceous C-type asteroids given their size, shape, spectra, albedo, rotational rate, and density. In fact, based on their similarity, one hypothesis is that both moons may be captured main-belt asteroids. They also appear to have a lunar-like uniform regolith layer that is tens of meters deep. Although infrared spectroscopy has not detected bound water in its surface minerals, both moons may be outgassing water vapor, as measured by magnetospheric data from the Phobos-2 spacecraft in 1989. Theoretical models indicate that if Phobos originally formed with an ice component, it could still retain an icy core at less than 100 m near the poles. Thus, the most likely composition of these moons is a combination of anhydrous silicates, carbon, organic compounds, and ice.

Phobos and Deimos hold a promising potential for volatiles resources because of their proximity to Mars, as well as being more accessible more often than any known asteroid. The amount of fuel necessary for an Earth-to-Phobos trip is less than that

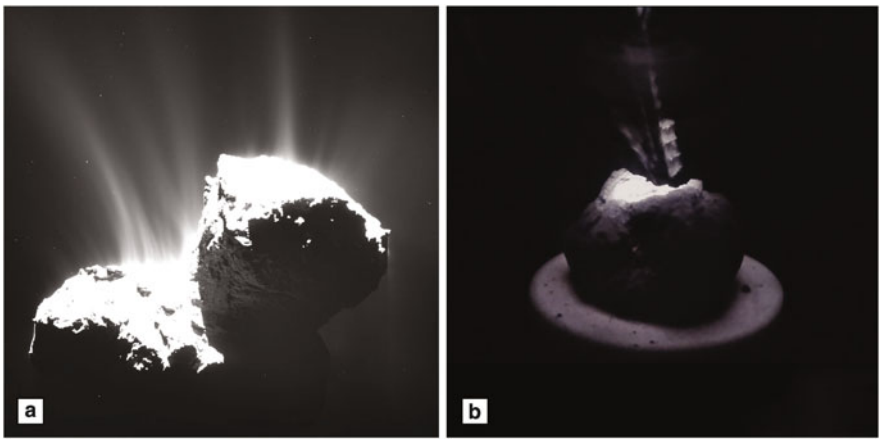

Fig. 15.4 (**a**) Image of comet 67P/Churyumov-Gerasimenko taken 84 miles from its center by ESA's Rosetta spacecraft showing jets of water and other volatiles spewing out from the comet; (**b**) image of a 2.5 cm sample of a carbonaceous chondrite meteorite heated by a focused light source and showing spalling and extraction of volatiles (water, CO_2, and SO_2) to test the effectiveness of optical mining for asteroids. Source: (**a**) ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/-IAA/SSO/INTA/UPM/DASP/IDA. (**b**) Colorado School of Mines, Sercel et al. (2016)

for a similar voyage to the lunar surface. Launch windows occur about every 2 years with traveling times of several months. As a result, they are considered convenient stepping stones for future human missions to Mars (Chap. 5). Establishing a base on these moons would be ideal to remotely study Mars, while avoiding biological contamination and the long communication delays from Earth. Several missions are in the plans, which should finally help assess their potential as a resource in the future.

15.3 Resource Mining

As shown in Table 15.1, a large variety and quantity of resources are present on the Moon, Mars, and NEAs. However, it is the combination of resource type, mining techniques, processing technology, and product demand that determines what resources are regarded as valuable, or specifically in the case of mineral deposits, as ore (an economic source of material). Once valuable deposits are identified, technologies need to be designed, developed, and tested to extract these materials. On Earth, we have been taking this approach for thousands of years making use of abundant energy, oxygen and water, a large industrial infrastructure and labor force, and adjusting all systems to our terrestrial gravitational force. However, none of these elements are similar in space. The lack of a thick atmosphere, gaseous oxygen, liquid water, a fully developed industrial base, and a human workforce in planetary bodies with low gravity levels make the process of space resource utilization a new and challenging task. Thus, it would make sense to first identify common valuable resources and then develop common processes and technologies to make efficient use of these resources. One should keep in mind that while this is the most effective and economical approach, it may not be possible to achieve it given the different conditions present on the various planetary destinations. From the inventory listed in Table 15.1, we can identify the common valuable resources that will help us in the near future to obtain the propellants, consumables, energy sources, and construction materials to allow us to live-off-the-land in space. Among the most precious materials are: water, carbon, metals, oxygen, silicon, volatiles, and raw regolith.

Just as in terrestrial mining practices, the main steps that should be taken to obtain and utilise these common resources in space are: (i) extraction, (ii) material handling and transport, and (iii) processing. This section provides a description of each step, its current status of development, and future challenges.

15.3.1 Extraction

After assessing resource availability, the process of extraction consists of separating the material from its original location, concentrating it, and converting it

Table 15.1 Overview of space resources

Moon	Mars	Asteroids	Potential use
Atmospheric and environmental resources			
Solar radiation	Solar radiation	Solar radiation	Power generation
Low gravity (1/6 g)	Low gravity (1/3 g)	Microgravity	Materials processing
Vacuum		Vacuum	Materials processing
	Wind		Power generation
	CO_2		Propellants, plastics
	Ar		Inert atmospheres
	H_2O		Life support, propellants
	N_2		Agriculture, inert gas
Raw unprocessed material resources			
Regolith	Regolith	Regolith	Construction
Lava tubes	Hills, craters	Craters	Radiation shielding
			Thermal protection
Non-metals			
(from oxides and silicates)	(from oxides, silicates, sulfates, phosphates)		
O	O		Life support, propellant
	S		Agriculture
Si	P, Si	P, Si, Ga, Ge, Se	Agriculture, photovoltaic cells, electronics
Volatiles			
(from cometary material trapped at poles)	(from polar ice, subsurface ice and hydrated minerals)		
H_2O, C, N	H_2O	H_2O, CO_2	Life support, propellants
	Deuterium		Fusion power
(from solar wind)			
H, C			Propellants
He, ^{3}He			Fusion power
N			Agriculture, inert gas
O			Life support, propellants
P, S			Agriculture
Metals			
(from silicates, oxides, meteorites)	(from oxides, carbonates, meteorites)		
Ti, Al, Fe, Ni, Mg	Fe, Ni, Al, Mg, Ti, Ca	Fe, Ni, Co	Fabrication of equipment, construction
(from KREEP)			
REE, K		REE, Pt group metals, Ag, Au	Electronics, photovoltaics, terrestrial import

Source: Lewis et al. (1993), Heiken et al. (1991), Kargel (1994), Eckart (2006), Crawford (2015), McKay et al. (1992)

for direct use, manufacturing, or further processing. This includes raw materials for production of propellants and life support gasses, habitats for humans and equipment, and fuel for power generation. Following these steps help identify if natural concentrators of elements exist (such as wind, water, or mantle uprising) or if mechanical and chemical processes are needed to obtain the desired elements.

Technologies for extracting mineral resources on Earth have been under development for thousands of years. Modern mining relies on drilling, fragmentation, excavation, and transportation of rock and soil, both underground and on the surface. Yet, the challenge is to effectively modify these terrestrially mature technologies to operate in lunar, asteroidal, and Martian conditions for acquiring and beneficiating regolith and rocks.

Understanding the processes that lead to the formation and segregation of natural materials in the various extraterrestrial locations is needed to learn how to identify resources from geological context. Equally important for extraction is predicting the behavior of regolith and granular materials on the Moon, Mars, and asteroids. Power, from imported sources or locally produced will have to be generated. Finally, telecommunications and navigation, robotic access to planetary surfaces, human exploration, and autonomous systems will need to be provided.

Extraction capabilities that have been demonstrated in space include scooping of regolith samples on the Moon and Mars, coring of regolith samples on the Moon, and grinding and analysis of rock samples on the Moon and Mars. Atmospheric gases have been acquired and separated on Mars with some basic filtration of dust. Ongoing research efforts have focused on developing technologies such as: (a) assessment, excavation, and processing of regolith from permanently shadowed lunar craters to release water, volatiles, and produce oxygen, (b) surface excavators capable of handling regolith and ice-regolith mixtures on the Moon and Mars, and (c) collection and purification of volatiles from moderate to high-vacuum environments.

As more emphasis is placed on mining the NEAs, adapting all these techniques to ultra-low gravity environments will be the greatest challenge. Given the little experience with handling materials in weightlessness, this will be an important effect to consider. Technologies for maneuvering excavation tools and containing the excavated material without the help of gravity will have to be developed. Given these challenges, a technique called *optical mining* is being investigated to drill and extract volatiles using focused solar radiation without having to dock or anchor any equipment to the surface of an asteroid (see Fig. 15.4b).

15.3.2 Material Handling and Transportation

The extraction and processing of space resources requires the handling and transportation of materials in a variety of temperatures, pressures, and gravity levels. The materials could be solid, liquid, vapor or multi-phase, which are affected by gravity and which may also require controlled containment. The transport of materials

may involve stationary devices such as augers and conveyors, surface vehicles with wheels, tracks, or rails, and flight vehicles including aircraft or hoppers.

As with extraction systems, material handling and transportation under terrestrial conditions has been done for thousands of years. The challenge lies on adapting such systems or designing new ones for other bodies in the Solar System. In the case of granular material, our knowledge is mostly empirical even on Earth and may not be scalable to other planetary surfaces given the different gravity levels, temperature, pressure and the abrasive and electrically charged nature of lunar, asteroidal, and martian regolith. Gravity also plays an important and poorly studied role on multi-phase and non-isothermal fluids, as well as on the interaction of hardware with regolith and soils. Such is the case of many types of mining equipment used for excavating, loading, and hauling, which are extremely dependent on adequate traction. The ability to exert large horizontal forces may be severely limited for this equipment under low-gravity conditions.

Experience on handling materials in other planetary bodies has been restricted to manipulation and encapsulation of lunar samples by the Apollo astronauts, who then transported these samples in small containers on rover vehicles. Martian surface samples have been robotically manipulated for analysis by the various exploration rovers. For the NASA Resource Prospector mission expected to be launched in the early 2020s, a mobile platform is currently being designed to drill and collect regolith to extract water and volatile components from the permanently shadowed lunar craters at temperatures as low as 40 K, hard vacuum, and 1/6th of Earth's gravity. Similar systems are also being developed for Mars to work under that planet's gravity, temperatures, and atmospheric conditions. From the various terrestrial handling and transportation systems analyzed, continuous digging machines (such as bucket wheel and bucket ladder excavators) have surfaced as one of the most efficient systems for excavating and transporting lunar and martian regolith (see Fig. 15.5). Emphasis is being placed on completely autonomous operation of equipment to avoid any human interaction. Finally, as a result of the abrasive, intrusive, and electrostatically charged nature of dust, mitigation techniques are being investigated to develop technologies to prevent and remove its accumulation and damaging effects on machinery, seals, and valves.

Asteroids pose their unique set of constraints and challenges. The lack of gravity makes terrestrial modes of transportation and sample handling impossible and techniques such anchoring, magnetic transport, and hopper hauling and docking should be explored. Several robotic sample-return and human missions to asteroids are being planned, which will have to address some of these techniques in the next few years.

15.3.3 Processing

The final steps in the chain of operations with space resources are the processing and conversion of raw materials into usable products. These products include fuel

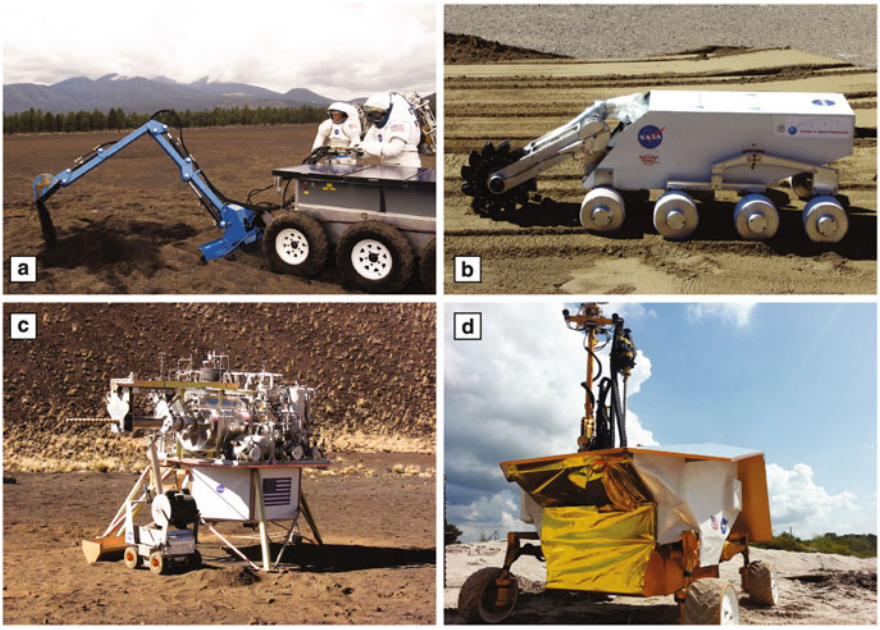

Fig. 15.5 Space mining equipment: (**a**) NASA's backhoe excavating a trench in a terrestrial analog field; (**b**) Colorado School of Mines bucket wheel excavator and transporter prototype; (**c**) Lockheed Martin's Precursor ISRU Lunar Oxygen Testbed (PILOT) system capable of producing oxygen by ilmenite reduction; (**d**) NASA's Resource Prospector mission prototype designed to drill and extract volatiles and water from a lunar polar crater. Source: (**a**, **d**) NASA. (**b**) Colorado School of Mines. (**c**) Courtesy Lockheed Martin

and oxygen for propulsion and mission consumables such as oxygen, water, and buffer gases for life support and fertilizer for plant growth, as well as materials for manufacturing parts and construction.

Since approximately 85% of the mass of any spacecraft launched from the surface of a planetary body is the fuel and oxidizer, propellant manufacturing in situ is an important product of resource processing. For missions involving humans, significant quantities of oxygen and water must be produced for life support, regardless of the use of closed-loop life support systems. Once permanent settlements in other planetary bodies are established, fertilizers to produce food will have to be processed. For extended missions on the lunar surface or when traveling to NEAs or Mars, the ability to manufacture parts on site will be critical. As shown by several ongoing research efforts, additive manufacturing (3D printing) could have a significant effect on the availability of tools and parts on the surface of the planet and also in orbiting spacecraft.

As is the case with extraction and handling of terrestrial resources, processing technologies and the chemical methods used have been known and applied for centuries. The challenge ahead is finding the way to implement these technologies

for extraterrestrial environments. For example, methods to produce oxygen from the regolith of the Moon, such as reduction of ilmenite with hydrogen or carbon monoxide, molten silicate electrolysis, and carbothermal reduction, have been achieved in laboratories for years and have even recently been tested in real-scale prototype processing plants on terrestrial analog sites. These oxygen production technologies leave behind pure metals as residue and laboratory experiments have been performed to separate these metals from the remaining slag. This slag can also prove useful as a feedstock for the production of bricks or construction blocks. For rocket fuel production, the chemical processes necessary to generate methane on the Moon with solar-wind implanted hydrogen and carbon are well known, although several methods are still under study to extract these elements from the regolith. A research project is currently being developed to determine the concentration and form of lunar polar hydrogen and to demonstrate the extraction of water ice from the regolith. This experiment is also expected to characterize the soil mechanics and the fine grain characteristics of the lunar polar soil and it is planned to be part of the NASA Resource Prospector mission.

For oxygen and fuel production on Mars, the focus has primarily been on atmospheric processing technologies. Several prototype systems for methane production have been developed based on the Sabatier method. Although oxygen is also generated as byproduct of the Sabatier reaction through the electrolysis of water, it is produced at very small quantities. Oxygen can instead be produced by solid oxide electrolysis, the reverse water-gas shift reaction, and cold plasma CO_2 dissociation. In fact, an experiment using solid-oxide electrolysis is slated to fly on the Mars 2020 lander, making it most probably the first ISRU technology demonstration in space. Laboratory demonstrations have also been performed to produce other hydrocarbon fuels, such as methanol, ethylene, and benzene.

For NEAs, water and carbon dioxide could be cold trapped after being extracted using the optical mining technique. The water could be directly used for steam propulsion, electrolyzed to obtain hydrogen and oxygen as rocket propellants, or reacted with carbon dioxide to produce methane fuel.

15.4 Resource Utilisation

More than 100,000 years ago our ancestors moved out of Africa driven by basic survival needs: they were looking for food and water after long periods of drought and population growth. In the process of migrating, they learned to generate and control fire, preserve food, and discovered agriculture, which allowed them to live in harsh climates, explore all continents, and eventually settle into increasingly complex communities. As early settlements gave way to towns, cities, and countries, and as these became more interdependent, the search for other resources became a priority: gold, silver, silk, and spices for trade, oil as the primary source of energy, advanced materials for construction and transportation, and more recently rare-earth elements that are now driving the renewable energy, telecommunications,

and electronics industries. It is clear that our incessant search for resources has been responsible for all terrestrial exploration, shaping of countries boundaries, and spurring rapid technological development. Resources are, in a few words, the main engine of civilization.

The quest for technological superiority and to control the higher ground for military purposes propelled the two main superpowers of the twentieth century, the United States and the Soviet Union, to explore beyond Earth's atmosphere. An unintended consequence of seeing our planet from outside for the first time was the realization of the great potential of utilizing space for our terrestrial advantage. In that way, Earth observation from orbit for weather forecasting, communications, and studies of climate and geological change, became the first space resource to be utilized.

However, after more than 50 years of space exploration, all robotic probes sent to destinations in the Solar System and all human missions to LEO and the Moon have, with the exception of solar energy, relied exclusively on resources brought from Earth for propulsion and life support. For several decades now, various groups inside and outside government space agencies have been advocating for the use of extraterrestrial resources to explore further, longer, and larger to increase scientific returns. In addition, starting a few years ago with the rise of private space companies, a call has also come to utilize resources for developing commercial ventures in space.

As we try to expand exploration further and extend the life and reach of interplanetary probes, orbiting satellites, and surface rovers, propellants extracted from the Moon, asteroids and planets will be needed to establish a sustainable and reliable space transportation infrastructure. Consequently, the use of space resources with the greatest priority and economic significance for both human and robotic missions may be the production of rocket propellants. If these propellants could be generated from local materials, either the mass of payload delivered to the planetary body could be significantly increased, or the size of the vehicle departing from Earth could be substantially reduced. For example, if hydrogen and oxygen are to be used as propellants, even if the hydrogen had to be carried along from Earth, the use of oxygen extracted in space could reduce the amount of propellant that must be lifted from Earth by a factor of ten. If oxygen can be produced reliably in situ, it also becomes feasible to operate planetary landers for both ascent and descent to rendezvous in orbit with a vehicle sent from Earth. In time, it may even be possible to export locally produced hydrogen and oxygen to the various Lagrangian points between the Earth, the Moon and Mars, and to LEO and GEO, where it might be economically attractive to use them as propellants in support of lunar resupply operations, missions to asteroids and Mars, and to refuel orbiting satellites. In fact, in the last few years several private companies have developed a variety of designs for fuel depots intended to be parked in LEO and Langrangian points, with the intention to commercially supply propellants.

As described in the previous section, the most accessible source for propellants appears to be volatiles, such as hydrogen, carbon, and water, which can be extracted by surface heating up to 700 °C. Water can then be electrolyzed to obtain oxygen

and additional hydrogen. On the Moon, water can be extracted from permanently shadowed craters while solar-wind implanted hydrogen and carbon may be obtained from the lunar regolith. If access to these deep polar craters proves to be difficult, oxygen could also obtained from minerals found on the lunar regolith away from the poles by ilmenite reduction or magma electrolysis. On Mars, water ice is readily available in the poles and also located below the surface or in hydrated minerals in other regions of the planet. Oxygen can also be easily extracted from the abundant carbon dioxide in the atmosphere. Mars is a volatile-rich planet, containing abundant quantities of light elements compared to our Moon. Fuels such as hydrogen, methane, methanol, ethylene, and hydrogen peroxide can all be made from these volatile elements.

As for human exploration, just as our ancestors did in prehistoric times, the earliest use of space resources will be for habitation purposes. Caves on lava tubes on the Moon and dwellings dug into the side of volcanic tuff yardangs on Mars may very well be our first homes in other celestial bodies. Raw lunar or Martian regolith could be used for radiation shielding, protection against micrometeoroid impacts and blowing dust raised by ascent and descent rocket engines, and for thermal insulation and heat storage. Next in complexity will be the search, extraction, and use of water. The vital liquid could be obtained by the methods described previously and may be used for human life support, growing crops, controlling dust in air-locks and inside habitats, as an additive for construction, and for protection against cosmic radiation. Interestingly, a given mass of water is a much better shield against energetic protons than an equal mass of lunar or martian rock, dirt, or even metals. Water can be pumped, used inside walls, or to cover the roofs of habitats. For growing plants for food and as a psychological benefit to astronauts, volatiles such as nitrogen could be used along with indigenous phosphorus for agricultural purposes. It would also be possible to raise crops using the abundant CO_2 in the atmosphere of Mars and water extracted from the soil. Breathing oxygen could be extracted from these substances, or as a byproduct of plant photosynthesis.

Future uses of extraterrestrial materials will most probably include the fabrication of metal components and structures, given that metals are byproducts from oxygen production from regolith. For example, metals can be used in the construction of electrical distribution systems on the Moon, both for local use and for solar-produced electric power which may be exported to Earth. The base of silicon solar cells can be manufactured from lunar soil and the trace elements needed to turn them into electric cells can be brought from Earth. Once arrays of large solar collectors are placed on several locations, energy can be beamed by laser or microwaves to other parts of the Moon and to rectennas located in various places on Earth. This scheme has been proposed by several studies in the past as one of the main rationales for going back to the Moon. In addition, metals found in abundance on Mars, such as magnesium or aluminum, have been proposed as fuel or propellant sources by reacting with oxygen or directly with atmospheric CO_2 or CO. Although not as efficient as metal-oxygen propellants, a metal-CO_2/CO rocket engine can be used for short-distance hoppers or even for ascent vehicles to Martian orbit. Finally, recycled parts from spent spacecraft are also a viable source for useful metals, while

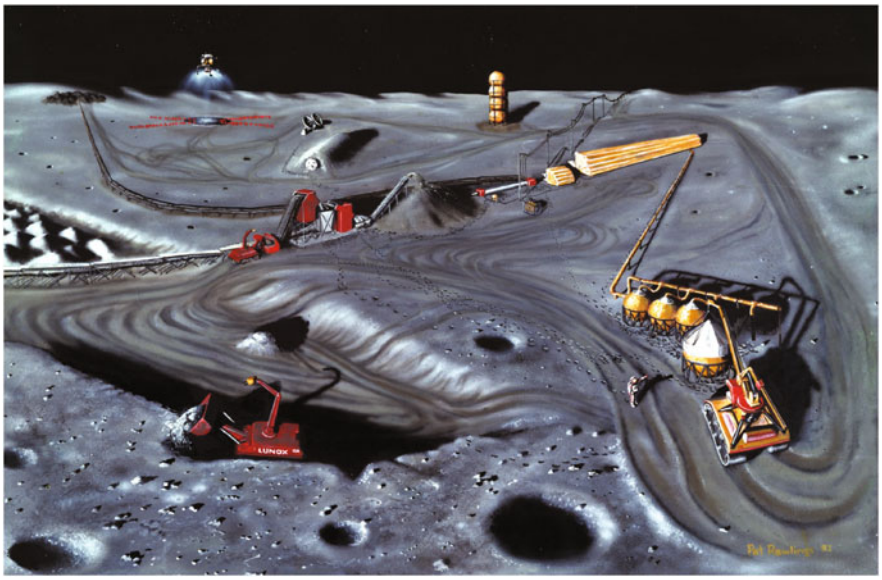

Fig. 15.6 Artist concept of a human settlement and surface mining operations on the Moon showing excavation, transportation, material handling, and processing equipment, as well as regolith covered habitats, roads, berms, launch pads, power lines, fuel depots, and communication antennas. Source: artwork by Pat Rawlings, painted in 1983

additive manufacturing could be utilised for replacement of parts and for large civil engineering projects (Fig. 15.6).

In the long term, resources from space brought to Earth could be used for terrestrial applications. These resources will have to be either of extreme value regardless of an increase in their supply, more economical to extract and transport from space, or prevent severe environmental damage on Earth by extracting them somewhere else. One such case advocated by space mining companies is the extraction of platinum group metals from asteroids that could prove to be of extreme value on Earth given their current high prices. However, finding easy-to-reach NEAs, identifying their composition, and developing economically and technically viable extraction techniques are all areas still in their infancy. In addition to beamed solar energy from the Moon to Earth, ^{3}He (helium-3) from the Moon and deuterium from Mars have been identified to be potential exports, which could be of value as fuel if large-scale fusion power generation ever becomes a reality on Earth. In the case of helium-3, although only present in a few ppb on the Moon, this concentration is orders of magnitude greater than any source of this isotope on Earth. However, extracting these minute amounts of helium-3 from the lunar regolith would require an extremely complex mining operation over a vast area. Unless more efficient extraction techniques are developed and fusion technology using helium-3 as fuel is not only achieved, but also used to generate electric power at an industrial level, mining and transportation of Helium-3 to Earth will remain an elusive objective.

Extraterrestrial resources may help both science and private companies advance their exploration objectives. However, whereas the former sees exploration as a way to gather data to better understand the universe in which we live, for the latter, exploration is an investigation for the purpose of locating opportunities for profit. Nevertheless, progress from one group can benefit the other's objectives if proper communication is established among them to make the best use of these resources. Finally, we have to consider the legal ramifications of exploiting resources in space. Contrary to the arbitrary claims and sometimes violent acquisition and exploitation of territories for resource extraction during most of human history, the international treaties that were crafted at the beginning of the space age were intended to ensure the peaceful use of space. The Outer Space Treaty of 1967 and the Moon Treaty of 1979 included language to prevent countries or individuals to claim ownership of any celestial body beyond Earth. However the language of these documents leave the definition of space resource mining, extraction, and utilization open for interpretation. As private companies have started working towards exploiting the resources from asteroids, the Moon, and Mars, an international consensus is needed to regulate commercial exploration and utilization of resources from planetary bodies. With the precedent set by centuries of chaotic and unregulated use of terrestrial resources, the time has come to start working towards an agreement at the international level to define fair and just terms to utilize space.

Space resources hold the secret to unleash an unprecedented wave of exploration by utilizing the full potential and value of space. Creating a space infrastructure based solely on space resources will allow us to cut the umbilical cord from Earth and fully incorporate space into our lives by expanding the frontiers of our ecosystem. This promises to increase our knowledge of the Solar System in unimaginable ways, while giving us access to abundant material wealth for millennia to come. It is up to us humans here on planet Earth to find the best way to utilize these extraterrestrial resources effectively and responsibly to make this promise a reality.

Take-Home Messages

Valuable extraterrestrial resources such as solar energy, low-gravity, ultra-high vacuum, water, oxygen, silicon, volatiles, metals, and minerals are available in outer space, planets, the Moon, asteroids, and comets.

Just like resources on Earth, space resources will enable exploration by providing the necessary energy, propellants, materials, and basic consumables to significantly reduce the mass, cost, and risk of space missions, while also extending their reach and duration.

Utilization of space resources will require their accurate identification by remote sensing and surface characterization, as well as their extraction, transportation, and processing under conditions very different from the ones on Earth.

Space resources will enable the development of a transportation and exploration infrastructure in space that will help further our understanding of the Solar System, while also helping to generate new commercial ventures beyond Earth.

Scientific goals and commercial interests will have to be properly balanced and guided by an international agreement to make the most effective use of extraterrestrial resources for the benefit of all humanity.

Suggested Readings

Abbud-Madrid, A., Modak, A., Branch, M.C., Daily, J.W.: Combustion of magnesium with carbon dioxide and carbon monoxide at low gravity. J. Propuls. Power **17**, 852–859 (2001)

Crawford, I.A.: Lunar resources: a review. Prog. Phys. Geogr. **39**, 137–167 (2015)

Duke, M.B., Gaddis, L.R., Taylor, G.J., Schmitt, H.H.: Development of the Moon, vol. 60, pp. 597–656. (2006) doi:10.2138/rmg.2006.60.6

Eckart, P. (ed.): The Lunar Base Handbook: An Introduction to Lunar Base Design, Development, and Operations. Space Technology Series, 2nd edn., p. 850. McGraw-Hill, New York (2006)

Feuerbacher, B., Stower., H. (eds.): Utilization of Space: Today and Tomorrow, p. 410. Springer, Berlin (2006)

Finney, B.R., Jones, E.M. (eds.): Interstellar Migration and the Human Experience. p. 354. University of California Press, Berkeley (1985)

Gaffey, M.J., Burbine, T.H., Binzel, R.P.: Asteroid spectroscopy: progress and perspectives. Meteoritics **28**, 161–187 (1993)

Haskin, L.A.: Toward a spartan scenario for use of lunar materials. In: Wendell, W.W. (ed.) Lunar Bases and Space Activities of the 21st century, pp. 435–443. Lunar and Planetary Institute, Houston (1985)

Heiken, G., Vaniman, D., French, B.M. (eds.): The Lunar Sourcebook: A User's Guide to the Moon, p. 736. Cambridge University Press, Cambridge (1991)

Jolliff, B.L., Wieczorek, M.A., Shearer, C.K., Neal, C.R. (eds.): New Views of the Moon. Mineralogical Society of America, Chantilly (2006)

Kargel, J.S.: Metalliferous asteroids as potential sources of precious metals. J. Geophys. Res. **99**, 21129–21141 (1994). doi:10.1029/94JE02141

Lewis, J., Matthews, M.S., Guerrieri, M.L. (eds.): Resources of Near Earth Space, p. 977. The University of Arizona Press, Tucson (1993)

Lewis, J.S.: Mining the Sky: Untold Riches from the Asteroids, Comets and Planets, p. 263. Helix Books, Reading (1996)

McKay, M.F., McKay, D.S., Duke, M.B. (eds.): Space Resources, vols. 1–4, No. 509 in NASA Special Publications. NASA, Houston (1992)

O'Neill, G.K., J., B., Gilbreath, W., O'Leary, B. (eds.): Space Resources and Space Settlements. No. 428 in NASA Special Publications, NASA Ames Research Center, Moffett Field (1979)

Reimold, W.U., Koeberl, C., Gibson, R.L., Dressler, B.O.: Economic mineral deposits in impact structures: a review. In: Impact Tectonics, pp. 479–552, Springer, Heidelberg (2005) doi:10.1007/3-540-27548-7_20

United Nations: United Nations Treaties and Principles on Outer Space. ST/SPACE/11/Rev.2, p. 80. United Nations Publications, New York (2009)

Zacharias, M., Gertsch, L., Abbud-Madrid, A., Blair, B., Zacny, K.: Real-world mining feasibility studies applied to asteroids, the Moon, and Mars. In: AIAA SPACE 2011 Conference & Exposition, AIAA-2011-7115 (2011)

Image and Data Sources

Amante, C., Eakins, B.: ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. Technical Report, National Geophysical Data Center, NOAA (2009)
Arndt, N.T., Nisbet, E.G.: Processes on the young Earth and the habitats of early life. Annu. Rev. Earth Planet. Sci. **40**, 521–549 (2012)
Arvidson, R.E., Bonitz, R.G., Robinson, M.L., Carsten, J.L., Volpe, R.A., Trebi-Ollennu, A., Shaw, A.S.: Results from the Mars Phoenix lander robotic arm experiment. J. Geophys. Res. **114**(E1), 21 (2009)
Basilevsky, A.T., Head, J.W.: The geologic history of Venus: a stratigraphic view. J. Geophys. Res. Planets **103**(E4), 8531–8544 (1998)
Bird, P.: An updated digital model of plate boundaries [digitized by Hugo Ahlenius, Nordpil and Peter Bird]. Geochem. Geophys. Geosyst. **4**(3), 1027 (2003), updated version from https://github.com/fraxen/tectonicplates
Brož, P., Čadek, O., Hauber, E., Rossi, A.P.: Shape of scoria cones on Mars: insights from numerical modeling of ballistic pathways. Earth Planet. Sci. Lett. **406**, 14–23 (2014)
Carr, M.H., Channels and valleys on Mars: cold climate features formed as a result of a thickening cryosphere. Planet. Space Sci. **44**(11), 1411–1423 (1996)
Carr, M.H., Head, J.W.: Geologic history of Mars. Earth Planet. Sci. Lett. **294**(3), 185–203 (2010)
Carter, J., Poulet, F., Bibring, J.-P., Mangold, N., Murchie, S.: Hydrous minerals on Mars as seen by the CRISM and OMEGA imaging spectrometers: Updated global view. J. Geophys. Res. Planets **118**(4), 831–858 (2013)
Crawford, I.A.: Lunar resources: a review. Prog. Phys. Geogr. **39**, 137–167 (2015). doi:10.1177/0309133314567585
de Kock, M.O., Evans, D.A., Beukes, N.J.: Validating the existence of Vaalbara in the Neoarchean. Precambrian Res. **174**(1), 145–154 (2009)
Di Achille, G., Hynek, B.M.: Ancient ocean on Mars supported by global distribution of deltas and valleys. Nat. Geosci. **3**(7), 459–463 (2010)
E2E-iSAG: Planning for Mars returned sample science: final report of the MSR End-to-End International Science Analysis Group (E2E-iSAG). Astrobiology **12**(3), 175–230 (2012)
Eckart, P. (ed.): The Lunar Base Handbook: An Introduction to Lunar Base Design, Development, and Operations. Space Technology Series, p. 850. McGraw-Hill, New York (2006)
Ehlmann, B.L., Edwards, C.S.: Mineralogy of the Martian surface. Annu. Rev. Earth Planet. Sci. **42**(1), 291–315 (2014)
Ehlmann, B.L., Mustard, J.F., Murchie, S.L., Bibring, J.-P., Meunier, A., Fraeman, A.A., Langevin, Y.: Subsurface water and clay mineral formation during the early history of Mars. Nature **479**(7371), 53–60 (2011)
Fassett, C.I., Head, J.W.: (2008a). The timing of Martian valley network activity: constraints from buffered crater counting. Icarus **195**(1), 61–89
Fassett, C.I., Head, J.W.: (2008b). Valley network-fed, open-basin lakes on Mars: distribution and implications for Noachian surface and subsurface hydrology. Icarus, **198**(1), 37–56
Fassett, C.I., Head, J.W., Baker, D.M.H., Zuber, M.T., Smith, D.E., Neumann, G.A., Solomon, S.C., Klimczak, C., Strom, R.G., Chapman, C.R., Prockter, L.M., Phillips, R.J., Oberst, J., Preusker, F.: Large impact basins on Mercury: Global distribution, characteristics, and modification history from MESSENGER orbital data. J. Geophys. Res. Planets **117**(E12), e00L08 (2012)
Florensky, C.P., Ronca, L.B., Basilevsky, A.T., Burba, G.A., Nikolaeva, O.V., Pronin, A.A., Trakhtman, A.M., Volkov, V.P., Zazetsky, V.V.: The surface of Venus as revealed by Soviet Venera 9 and 10. Geol. Soc. Am. Bull. **88**(11), 1537–1545 (1977)
Fortezzo, C., Hare, T.: Completed digital renovation of the 1: 5,000,000 lunar geologic map series. In: Lunar and Planetary Science Conference, vol. 44, p. 2114 (2013)
Galilei, G.: Sidereus nuncius. Thomas Baglioni, Venezia (1610)

Geiss, J., Rossi, A.P.: On the chronology of lunar origin and evolution. Astron. Astrophys. Rev. **21**(1), 1–54 (2013)
Gilbert, G.K.: Sidereus nuncius. Philosophical Society of Washington, Washington (1893)
Goffredo, S., Dubinsky, Z. (eds.): Emergence and evolution of early life in the geological environment. The Cnidaria, Past, Present and Future, pp. 3–13. Springer, Berlin (2016)
Grant, J.A., Irwin, R.P., Grotzinger, J.P., Milliken, R.E., Tornabene, L.L., McEwen, A.S., Weitz, C.M., Squyres, S.W., Glotch, T.D., Thomson, B.J.: HiRISE imaging of impact megabreccia and sub-meter aqueous strata in Holden Crater, Mars. Geology **36**(3), 195–198 (2008)
Green, J., Short, N.M.: Volcanic Landforms and Surface Features: A Photographic Atlas and Glossary. Springer, New York (1971)
Hackman, R.J., Mason, A.C.: Engineer Special Study of the Surface of the Moon. US Army Map Service (1960)
Hare, T.M., Hayward, R.K., Blue, J.S., Archinal, B.A., Robinson, M.S., Speyerer, E.J., Wagner, R.V., Smith, D.E., Zuber, M.T., Neumann, G.A., Mazarico, E.: Image mosaic and topographic map of the Moon. Scientific Investigations Map, vol. 3316. U.S. Geological Survey, Reston (2015) URL http://pubs.usgs.gov/sim/3316/.
Hartmann, W., Raper, O.: The New Mars - The Discoveries of Mariner 9. NASA SP-337. NASA, Washington, D.C. (1974)
Hauber, E., Kronberg, P.: Tempe Fossae, Mars: A planetary analogon to a terrestrial continental rift? J. Geophys. Res. Planets **106**(E9), 20587–20602 (2001)
Hauber, E., Kronberg, P.: The large Thaumasia graben on Mars: is it a rift? J. Geophys. Res. Planets **110**(E7), e07003 (2005)
Hayward, R., Fenton, L., Tanaka, K., Titus, T., Colaprete, A., Christensen, P.: Mars global digital dune database: MC-1. Technical Report, US Geological Survey (2010)
Hayward, R., Fenton, L., Titus, T., Colaprete, A., Christensen, P.: Mars global digital dune database: MC-30. Technical Report, US Geological Survey, Reston (2012)
Hayward, R.K., Mullins, K.F., Fenton, L., Hare, T., Titus, T., Bourke, M., Colaprete, A., Christensen, P. (2007). Mars global digital dune database: MC2-MC29. Technical Report, Geological Survey (US).
Head, J.W.: Surfaces and interiors of the terrestrial planets. In: Beatty, J.K., Petersen, C.C., Chaikin, A. (eds.) The New Solar System. Cambridge University Press, Cambridge (1999)
Head, J.W., Chapman, C.R., Domingue, D.L., Hawkins III, S.E., McClintock, W.E., Murchie, S.L., Prockter, L.M., Robinson, M.S., Strom, R.G., Watters, T.R.: The geology of Mercury: the view prior to the MESSENGER mission. Space Sci. Rev. **131**(1-4): 41–84 (2007)
Heiken, G., Vaniman, D., French, B.M. (eds.): (1991a). The Lunar Sourcebook: A User's Guide to the Moon, p. 736. Cambridge University Press, Cambridge
Heiken, G.H., Vaniman, D.T., French, B.M. (eds.): (1991b). Lunar Sourcebook: A User's Guide to the Moon, p. 736. Cambridge University Press, Cambridge
Hiesinger, H., Head, J.W.: New views of lunar geoscience: An introduction and overview. Rev. Mineral. Geochem. **60**(1), 1–81 (2006)
Hoffman, P.F., Schrag, D.P.: The snowball Earth hypothesis: testing the limits of global change. Terra nova, **14**(3), 129–155 (2002)
Hörz, F., Grieve, R., Heiken, G., Spudis, P., Binder, A.: Lunar surface processes. In: Heiken, G.H., Vaniman, D.T., French, B.M. (eds.) Lunar Sourcebook: A User's Guide to the Moon, pp. 61–120. Cambridge University Press, Cambridge (1991)
Ivanov, M.A., Head, J.W.: Global geological map of Venus. Planet. Space Sci. **59**(13), 1559–1600 (2011)
Kargel, J.S.: Metalliferous asteroids as potential sources of precious metals. J. Geophys. Res. **99**, 21129–21141 (1994). doi:10.1029/94JE02141.
Kite, E.S., Howard, A.D., Lucas, A.S., Armstrong, J.C., Aharonson, O., Lamb, M.P.: Stratigraphy of Aeolis Dorsa, Mars: Stratigraphic context of the great river deposits. Icarus **253**, 223–242 (2015)

Kovesi, P.: Good Colour Maps: How to Design Them. CoRR (2015). abs/1509.03700, URL http://arxiv.org/abs/1509.03700

Kreslavsky, M.A., Ivanov, M.A., Head, J.W.: The resurfacing history of Venus: Constraints from buffered crater densities. Icarus **250**, 438–450 (2015)

Lewis, J., Matthews, M.S., Guerrieri, M.L. (eds.): Resources of Near Earth Space, p. 977. The University of Arizona Press, Tucson (1993)

Lucchitta, B.K.: Geologic map of the north side of the Moon. Scientific Investigations Map, US Geological Survey Reston, Reston (1978)

Lynch, D.K., Hudnut, K.W., Dearborn, D.S.: Low-altitude aerial color digital photographic survey of the San Andreas fault. Seismol. Res. Lett. **81**(3), 453–459 (2010)

Marchi, S., Chapman, C.R., Fassett, C.I., Head, J.W., Bottke, W.F., Strom, R.G.: Global resurfacing of Mercury 4.0–4.1 billion years ago by heavy bombardment and volcanism. Nature, **499**(7456), 59–61 (2013)

Massé, M., Bourgeois, O., Le Mouélic, S., Verpoorter, C., Spiga, A., Le Deit, L.: Wide distribution and glacial origin of polar gypsum on Mars. Earth Planet. Sci. Lett. **317**, 44–55 (2012)

Massironi, M., Simioni, E., Marzari, F., Cremonese, G., Giacomini, L., Pajola, M., Jorda, L., Naletto, G., Lowry, S., El-Maarry, M.R. et al.: Two independent and primitive envelopes of the bilobate nucleus of comet 67P. Nature **526**(7573), 402–405 (2015)

McKay, M.F., McKay, D.S., Duke, M.B. (eds.): Space Resources, Volumes 1–4, NASA Special Publication, vol. 509. National Aeronautics and Space Administration, Washington, DC (1992)

McLennan, S.M. et al.: Geochemistry of sedimentary processes on Mars. Sedimentary Geology of Mars, vol. 102, pp. 119–138. Society for Sedimentary Geology, Tulsa (2012)

Melosh, H.J.: Global tectonics of a despun planet. Icarus **31**(2), 221–243 (1977)

Mitrofanov, I.G.: Hydrogen mapping of the lunar south pole using the LRO neutron detector experiment LEND. Science **330**, 483–486 (2010)

Mueller, K., Golombek, M.: Compressional structures on Mars. Annu. Rev. Earth Planet. Sci. **32**, 435–464 (2004)

Müller, R.D., Sdrolias, M., Gaina, C., Roest, W.R.: Age, spreading rates, and spreading asymmetry of the world's ocean crust. Geochem. Geophys. Geosyst. **9**(4), Q04006 (2008)

Nance, R.D., Murphy, J.B., Santosh, M.: The supercontinent cycle: a retrospective essay. Gondwana Res. **25**(1), 4–29 (2014)

Neukum, G., Oberst, J., Hoffmann, H., Wagner, R., Ivanov, B.: Geologic evolution and cratering history of Mercury. Planet. Space Sci. **49**(14), 1507–1521 (2001)

Pajola, M., Vincent, J.-B., Güttler, C., Lee, J.-C., Bertini, I., Massironi, M., Simioni, E., Marzari, F., Giacomini, L., Lucchetti, A. et al.: Size-frequency distribution of boulders $\geq$7 m on comet 67P/Churyumov-Gerasimenko. Astron. Astrophys. **583**, A37 (2015)

Platz, T., Byrne, P., Massironi, M., Hiesinger, H.: Volcanism and tectonism across the inner solar system: an overview. Geol. Soc. Lond. Spec. Publ. **401**(1), 1–56 (2015)

Pondrelli, M., Baliva, A., Di Lorenzo, S., Marinangeli, L., Rossi, A.P.: Complex evolution of paleolacustrine systems on Mars: an example from the Holden crater. J. Geophys. Res. Planets **110**(E4), E04016 (2005)

Pondrelli, M., Rossi, A.P., Le Deit, L., Fueten, F., van Gasselt, S., Glamoclija, M., Cavalazzi, B., Hauber, E., Franchi, F., Pozzobon, R.: Equatorial layered deposits in Arabia Terra, Mars: facies and process variability. Geol. Soc. Am. Bull. **127**(7–8), 1064–1089 (2015)

Prockter, L.M., Kinczyk, M.J., Byrne, P.K., Denevi, B.W., Head III, J.W., Fassett, C.I., Whitten, J.L., Thomas, R.J., Buczkowski, D.L., Hynek, B.M., Ostrach, L.R., Blewett, D.T., Ernst, C.M., the MESSENGER mapping group: The first global geologic map of Mercury. In: Lunar and Planetary Science Conference, vol. 47, p. 1245 (2016)

Sautter, V., Toplis, M., Wiens, R., Cousin, A., Fabre, C., Gasnault, O., Maurice, S., Forni, O., Lasue, J., Ollila, A. et al.: In situ evidence for continental crust on early Mars. Nat. Geosci. **8**(8), 605–609 (2015a)

Sautter, V., Toplis, M., Wiens, R., Cousin, A., Fabre, C., Gasnault, O., Maurice, S., Forni, O., Lasue, J., Ollila, A. et al.: In situ evidence for continental crust on early Mars. Nat. Geosci. **8**(8), 605–609 (2015b)

Scott, D.H., McCauley, J.F.: Geologic map of the west side of the Moon. Scientific Investigations Map, USGS, Reston (1977)
Sercel, J., Dreyer, C., Abbud-Madrid, A., Britt, D., Jedicke, R., Gertsch, L., Love, S.: A coordinated research program to develop the technology to optically mine asteroids. In: Proceedings of the 2016 American Society of Civil Engineers (ASCE) Earth and Space Conference. American Society of Civil Engineers (ASCE), Reston (2016)
Shearer, C.K., Hess, P.C., Wieczorek, M.A., Pritchard, M.E., Parmentier, E.M., Borg, L.E., Longhi, J., Elkins-Tanton, L.T., Neal, C.R., Antonenko, I. et al.: Thermal and magmatic evolution of the Moon. Rev. Mineral. Geochem. **60**(1), 365–518 (2006)
Sigurdsson, H.: Volcanic episodes and rates of volcanism. Encyclopedia of Volcanoes, pp. 271–279. Academic Press, New York (2000)
Solomon, S.C.: On volcanism and thermal tectonics on one-plate planets. Geophys. Res. Lett. **5**(6), 461–464 (1978)
Spudis, P., Pieters, C.: Global and regional data about the Moon. Lunar Sourcebook: A User's Guide to the Moon, pp. 595–632. Cambridge University Press, Cambridge (1991)
Stuart-Alexander, D.E.: Geologic map of the central far side of the Moon. Scientific Investigations Map. USGS, Reston (1978)
Sun, J., Muhs, D.: Dune fields, mid-latitudes. In: Encyclopedia of quaternary science, pp. 607–626, Elsevier, Amsterdam (2007)
Tanaka, K., Skinner, J., Dohm, J., Irwin, R., Kolb, E., Fortezzo, C., Platz, T., Michael, G., Hare, T.: Geologic map of mars. Scientific Investigations Map, vol. 3292. U.S. Geological Survey, Reston (2014)
Taylor, S.R., McLennan, S.: Planetary Crusts: Their Composition, Origin and Evolution, vol. 10. Cambridge University Press, Cambridge (2009)
USGS: Topographic map of mars, vol. I-2782 of Scientific Investigations Map. U.S. Geological Survey, Reston (2003). URL http://pubs.usgs.gov/imap/i2782/.
van Gasselt, S.: Cold-Climate Landforms on Mars, p. 292. Shaker, Aachen (2007)
van Gasselt, S., Neukum, G.: Chronology, cratering and stratigraphy. In: Encyclopedia of Astrobiology, pp. 304–313, Springer, Berlin (2011)
Van Kranendonk, M., Altermann, W., Beard, B., Hoffman, P., Johnson, C., Kasting, J., Melezhik, V., Nutman, A., Papineau, D., Pirajno, F.: The Geologic Time Scale 2012. Springer, New York (2012)
Vincent, J.-B., Bodewits, D., Besse, S., Sierks, H., Barbieri, C., Lamy, P., Rodrigo, R., Koschny, D., Rickman, H., Keller, H.U. et al.: Large heterogeneities in comet 67P as revealed by active pits from sinkhole collapse. Nature, **523**(7558): 63–66 (2015)
Werner, S.C.: Moon, Mars, Mercury: basin formation ages and implications for the maximum surface age and the migration of gaseous planets. Earth Planet. Sci. Lett. **400**, 54–65 (2014)
Wilhelms, D., McCauley, J., Trask, N.: The Geologic History of the Moon. USGS, Reston (1987)
Wilhelms, D.E., El-Baz, F.: Geologic map of the east side of the Moon. Scientific Investigations Map, US Geological Survey, Reston (1977)
Wilhelms, D.E., Howard, K.A., Wilshire, H.G.: Geologic map of the south side of the Moon. Scientific Investigations Map, Department of the Interior. US Geological Survey, Reston (1979)
Wilhelms, D.E., McCauley, J.F.: Geologic map of the near side of the Moon. Scientific Investigations Map, US Geological Survey, Reston (1971)
Wilson, L., Head, J.W.: Mars: review and analysis of volcanic eruption theory and relationships to observed landforms. Rev. Geophys. **32**(3), 221–263 (1994)

Appendix: Planetary Facts, Data and Tools

Planetary Constants

See Tables A1 and A2.

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8

Table A1 Bulk parameters for planets, dwarf planets and selected satellites

	Discovery	Moons	Mass m [$\times 10^{24}$kg]	Polar radius r_p [km]	Equatorial radius r_e [km]	Inverse flattening $1/f$ [–]	Density ρ [kg/m^3]	Gravity g [m/s^2]	Magnetic field B [T]	Atmospheric pressure p [bar]
Planets										
Mercury	prehistoric	0	0.330	2439.7	2439.7	–	5427	3.710	3.0×10^{-7}	10^{-14}
Venus	prehistoric	0	4.868	6051.8	6051.8	–	5243	8.870	–	92
Earth	prehistoric	1	5.972	6356.8	6378.1	298.253	5514	9.810	2.4×10^{-5}	1.014
Mars	prehistoric	2	0.642	3376.2	3396.2	169.894	3933	3.710	–	0.006
Jupiter	prehistoric	67	1898.190	66854.0	71492.0	15.41	1326	24.790	4.3×10^{-4}	> 1000
Saturn	prehistoric	62	568.340	54364.0	60268.0	10.21	687	10.440	2.2×10^{-5}	> 1000
Uranus	1781	27	86.813	24973.0	25559.0	43.62	1271	8.870	2.3×10^{-5}	> 1000
Neptune	1846	14	102.413	24341.0	24764.0	58.54	1638	11.150	1.4×10^{-5}	> 1000
Dwarf planets			m [$\times 10^{21}$ kg]							
(134340) Pluto	1930	5	13.030	1187.0	1187.0	–	1860	0.620	–	1×10^{-9}
(1) Ceres	1801	0	0.939	473.0	473.0	–	2161	0.280	–	–
(136199) Eris	2005	1	16.600	1163.0	1163.0	–	2520	0.820	–	–
(136472) Makemake	2005	1	< 4.400	715.0– 739.0	715.0– 739.0	20.00	1400– 3200	0.500	–	4–12×10^{-9}
(136108) Haumea	2004	2	4.010	~ 620.0	~ 620.0	–	(min) 2600	0.630	–	–

Satellites			m [$\times 10^{21}$ kg]							
Earth's moon	prehistoric		73.456	1736.0	1738.1	827.67	3344	1.620	1.0×10^{-7}	1×10^{-7}
Phobos (Mars I)	1877		10.659×10^{15}	11.3	11.3	–	1876	0.006	–	–
Deimos (Mars II)	1877		1.476×10^{15}	6.2	6.2	–	1471	0.003	–	–
Io (Jupiter I)	1610		89.319	1821.6	1821.6	–	3528	1.796	–	–
Europa (Jupiter II)	1610		47.998	1560.8	1560.8	–	3013	1.314	1.2×10^{-7}	10^{-12}
Ganymede (Jupiter III)	1610		148.190	2634.1	2634.1	–	1936	1.428	1.20×10^{-7}	–
Callisto (Jupiter IV)	1610		107.598	2410.3	2410.3	–	1834	1.235		7.5×10^{-12}
Mimas (Saturn I)	1789		0.037	198.2	198.2	–	1148	0.064	–	–
Enceladus (Saturn II)	1789		0.108	252.1	252.1	–	1609	0.113	–	trace
Tethys (Saturn III)	1684		0.617	531.1	531.1	–	984	0.146	–	–
Dione (Saturn IV)	1684		1.095	561.4	561.4	–	1478	0.232	–	–
Rhea (Saturn V)	1672		2.307	763.8	763.8	–	1.236	0.264	–	–
Titan (Saturn VI)	1655		134.500	2575.0	2575.0	–	1880	1.350	–	1.47
Iapetus (Saturn VIII)	1671		1.806	734.5	734.5	–	1088	0.223	–	–

Table A2 Orbital and axis parameters for planets, dwarf planets and selected satellites

	Orbit period T [d]	Periapsis q [$\times 10^6$ km]	Apoapsis Q [$\times 10^6$ km]	Semi major axis a [$\times 10^6$ km]	Orbit eccentricity [–]	Rotation period P [h]	Orbit inclination i [°]	Axis obliquity ε [°]
Planets								
Mercury	88.0	46.00	69.82	57.91	0.206	1407.60	7.00	0.03
Venus	224.7	107.48	108.94	108.21	0.007	−5832.60	2.64	177.36
Earth	365.3	147.09	152.10	149.60	0.017	23.93	0.00	23.44
Mars	687.0	206.62	249.23	227.93	0.094	24.62	1.85	25.19
Jupiter	4332.6	740.52	816.62	778.57	0.049	9.93	1.30	3.13
Saturn	10759.2	1352.55	1514.50	1433.53	0.057	10.66	2.49	26.73
Uranus	30685.4	2741.30	3003.62	2872.46	0.046	−17.24	0.77	97.77
Neptune	60189.0	4444.45	4545.67	4495.06	0.011	16.11	1.77	28.32
Dwarf planets								
(134340) Pluto	90.6	4436.82	7375.93	5906.38	0.249	−153.29	17.16	57.47
(1) Ceres	1681.63	382.62	445.41	414.02	0.076	9.07	10.59	4.00
(136199) Eris	203830.0	5723.00	14602.00	10162.50	0.441	25.90	44.04	?
(136472) Makemake	112897.0	5772.98	7904.75	6838.87	0.156	7.77	29.01	?
(136108) Haumea	103774.0	5228.74	7701.75	6465.25	0.191	3.92	28.19	?

Satellites	q [$\times 10^3$ km]		a [$\times 10^3$ km]		To plane			
Earth's moon	27.32	362.60	405.40	384.00	0.055	27.32	5.15	6.69
Phobos (Mars I)	0.32	9.23	9.52	9.38	0.015	synchronous	1.09	0.00
Deimos (Mars II)	1.26	23.46	23.47	23.46	0.000	synchronous	0.93	0.00
Io (Jupiter I)	1.77	420.00	423.40	421.70	0.004	synchronous	0.05	
Europa (Jupiter II)	3.55	664.86	676.94	670.90	0.009	synchronous	0.47	0.10
Ganymede (Jupiter III)	7.15	1069.20	1071.60	1070.40	0.001	synchronous	0.20	0.33
Callisto (Jupiter IV)	16.69	1869.00	1897.00	1883.00	0.007	synchronous	0.19	0.00
Mimas (Saturn I)	0.94	181.90	189.18	185.54	0.020	synchronous	1.57	0.00
Enceladus (Saturn II)	1.37	236.92	239.16	238.04	0.005	synchronous	0.02	0.00
Tethys (Saturn III)	1.89	294.62	294.62	294.62	0.000	synchronous	1.12	0.00
Dione (Saturn IV)	2.74	376.57	378.23	377.30	0.002	synchronous	0.02	0.00
Rhea (Saturn V)	4.52	526.51	527.57	527.04	0.001	synchronous	0.35	0.00
Titan (Saturn VI)	15.95	1186.15	1257.51	1221.83	0.029	synchronous	0.33	1.94
Iapetus (Saturn VIII)	79.32	3460.60	3662.00	3561.30	0.028	synchronous	15.47	0.00

Planetary Exploration Missions

A complete list of planetary exploration missions is provided in Table A3. Most of those missions are focused on an individual target body (e.g. Mars), several are covering multiple ones (e.g. Cassini-Huygens to the Saturn system, or Clementine to the Moon and asteroid 1620 Geographos). In the majority of cases also disciplines other than Geology are covered by missions' science objectives and their experiments.

Data and Tools

Planetary Geology, with the notable exception of the study of Meteorites or returned samples by either robots or humans, is largely based on remotely collected data. Those data have historically been shared within large communities in a relatively open fashion. Even during the Cold War, cooperation was active across US and Soviet scientists involved in planetary exploration.

Nowadays, planetary data are hosted and curated in dedicated archives that make available to anyone a range of science data products: from raw to calibrated, derived data (often described as *higher-level data*) (Table A4, as well as outreach products based on those, e.g. NASA Planetary Photojournal.[1]

The Planetary Data System (PDS) stands both for (1) the standards used in archiving planetary data (used also beyond NASA, that first developed them), (2) the organisation responsible of distributing and preserving data according to those standards, as well as (3) the distributed archives physically hosting those data.

Please note that data provided in this appendix might have a lifetime shorter than that of a book. Most agency and government URLs are likely to be available indefinitely or suitably redirected, though.

Please refer, for an updated view, to the GitHubrepository.[2]

We also suggest to monitor resource collections, listed below, maintained by long-term archives, such as NASA PDSand ESA PSA or any other provider associated to the IPDA

Code for introductory data handling of planetary data is available on the book's companion free GitHub repository.

[1] http://photojournal.jpl.nasa.gov.

[2] https://github.com/openplanetary/planetarygeology-book.

Table A3 Planetary missions until the end of 2016 (source: NASA NSSDC)

Launch date	Nation	Mission name	Notes	Target
1959-01-02	USSR	Luna 1	Flyby	Moon
1959-03-03	USA	Pioneer 4	Flyby	Moon
1959-09-12	USSR	Luna 2	Impact	Moon
1959-10-04	USSR	Luna 3	Probe	Moon
1960-10-10	USSR	Marsnik 1	Mars Flyby (Failure)	Mars
1960-10-14	USSR	Marsnik 2	Mars Flyby (Failure)	Mar
1961-02-04	USSR	Sputnik 7	Venus Impact (Failure)	Venus
1961-02-12	USSR	Venera 1	Venus Flyby (Failure)	Venus
1961-08-23	USA	Ranger 1	Test Flight (Failure)	Moon
1961-11-18	USA	Ranger 2	Test Flight (Failure)	Moon
1962-01-26	USA	Ranger 3	Impact (Failure)	Moon
1962-04-23	USA	Ranger 4	Impact	Moon
1962-08-25	USSR	Sputnik 19	Venus Flyby (Failure)	Venus
1962-08-27	USA	Mariner 2	Venus Flyby	Venus
1962-09-01	USSR	Sputnik 20	Venus Flyby (Failure)	Venus
1962-09-12	USSR	Sputnik 21	Venus Flyby (Failure)	Venus
1962-10-18	USA	Ranger 5	Impact (Failure)	Moon
1962-10-24	USSR	Sputnik 22	Attempted Mars Flyby	Mars
1962-11-01	USSR	Mars 1	Mars Flyby (Failure)	Mars
1962-11-04	USSR	Sputnik 24	Attempted Mars Lander	Mars
1963-04-02	USSR	Luna 4	Flyby	Moon
1963-11-11	USSR	Cosmos 21	Test Flight (Failure)	Venus
1964-01-30	USA	Ranger 6	Impact	Moon
1964-02-19	USSR	Venera 1964A	Venus Flyby (Failure)	Venus
1964-03-01	USSR	Venera 1964B	Venus Flyby (Failure)	Venus
1964-03-27	USSR	Cosmos 27	Venus Flyby (Failure)	Venus
1964-04-02	USSR	Zond 1	Venus Flyby (Failure)	Venus
1964-07-28	USA	Ranger 7	Impact	Moon
1964-11-05	USA	Mariner 3	Attempted Mars Flyby	Mars
1964-11-28	USA	Mariner 4	Mars Flyby	Mars
1964-11-30	USSR	Zond 2	Mars Flyby (Contact Lost)	Mars
1965-02-17	USA	Ranger 8	Impact	Moon
1965-03-21	USA	Ranger 9	Impact	Moon
1965-05-09	USSR	Luna 5	Impact	Moon
1965-06-08	USSR	Luna 6	Attempted Lander	Moon
1965-07-18	USSR	Zond 3	Lunar Flyby—Mars Test Vehicle	Mars
1965-07-18	USSR	Zond 3	Flyby	Moon
1965-10-04	USSR	Luna 7	Impact	Moon
1965-11-12	USSR	Venera 2	Venus Flyby (Failure)	Venus

(continued)

Table A3 (continued)

Launch date	Nation	Mission name	Notes	Target
1965-11-16	USSR	Venera 3	Venus Lander (Failure)	Venus
1965-11-23	USSR	Cosmos 96	Attempted Venus Lander?	Venus
1965-11-23	USSR	Venera 1965A	Venus Flyby (Failure)	Venus
1965-12-03	USSR	Luna 8	Impact	Moon
1966-01-31	USSR	Luna 9	Lander	Moon
1966-03-31	USSR	Luna 10	Orbiter	Moon
1966-05-30	USA	Surveyor 1	Lander	Moon
1966-08-10	USA	Lunar Orbiter 1	Orbiter	Moon
1966-08-24	USSR	Luna 11	Orbiter	Moon
1966-09-20	USA	Surveyor 2	Lander (Failure)	Moon
1966-10-22	USSR	Luna 12	Orbiter	Moon
1966-11-06	USA	Lunar Orbiter 2	Orbiter	Moon
1966-12-21	USSR	Luna 13	Lander	Moon
1967-02-04	USA	Lunar Orbiter 3	Orbiter	Moon
1967-04-17	USA	Surveyor 3	Lander	Moon
1967-05-08	USA	Lunar Orbiter 4	Orbiter	Moon
1967-06-12	USSR	Venera 4	Venus Probe	Venus
1967-06-14	USA	Mariner 5	Venus Flyby	Venus
1967-06-17	USSR	Cosmos 167	Venus Probe (Failure)	Venus
1967-07-14	USA	Surveyor 4	Lander (Failure)	Moon
1967-07-19	USA	Explorer 35	Orbiter	Moon
1967-08-01	USA	Lunar Orbiter 5	Orbiter	Moon
1967-09-08	USA	Surveyor 5	Lander	Moon
1967-11-07	USA	Surveyor 6	Lander	Moon
1968-01-07	USA	Surveyor 7	Lander	Moon
1968-04-07	USSR	Luna 14	Orbiter	Moon
1968-09-15	USSR	Zond 5	Return Probe	Moon
1968-11-10	USSR	Zond 6	Return Probe	Moon
1968-12-21	USA	Apollo 8	Crewed Orbiter	Moon
1969-01-05	USSR	Venera 5	Venus Probe	Venus
1969-01-10	USSR	Venera 6	Venus Probe	Venus
1969-02-25	USA	Mariner 6	Mars Flyby	Mars
1969-03-27	USA	Mariner 7	Mars Flyby	Mars
1969-03-27	USSR	Mars 1969A	Mars Orbiter (Failure)	Mars
1969-04-02	USSR	Mars 1969B	Mars Orbiter (Failure)	Mars
1969-05-18	USA	Apollo 10	Orbiter	Moon
1969-07-13	USSR	Luna 15	Orbiter	Moon
1969-07-16	USA	Apollo 11	Crewed Landing	Moon
1969-08-07	USSR	Zond 7	Return Probe	Moon
1969-11-14	USA	Apollo 12	Crewed Landing	Moon
1970-04-11	USA	Apollo 13	Crewed Landing (aborted)	Moon

(continued)

Table A3 (continued)

Launch date	Nation	Mission name	Notes	Target
1970-08-17	USSR	Venera 7	Venus Lander	Venus
1970-08-22	USSR	Cosmos 359	Attempted Venus Probe	Venus
1970-09-12	USSR	Luna 16	Sample Return	Moon
1970-10-20	USSR	Zond 8	Return Probe	Moon
1970-11-10	USSR	Luna 17	Rover	Moon
1971-01-31	USA	Apollo 14	Crewed Landing	Moon
1971-05-09	USA	Mariner 8	Mars Flyby (Failure)	Mars
1971-05-10	USSR	Cosmos 419	Attempted Mars Orbiter/Lander	Mars
1971-05-19	USSR	Mars 2	Mars Orbiter/ Attempted Lander	Mars
1971-05-28	USSR	Mars 3	Mars Orbiter/ Lander	Mars
1971-05-30	USA	Mariner 9	Mars Orbiter	Mars
1971-07-26	USA	Apollo 15	Crewed Landing	Moon
1971-09-02	USSR	Luna 18	Impact	Moon
1971-09-28	USSR	Luna 19	Orbiter	Moon
1972-02-14	USSR	Luna 20	Sample Return	Moon
1972-03-27	USSR	Venera 8	Venus Probe	Venus
1972-03-31	USSR	Cosmos 482	Attempted Venus Probe	Venus
1972-04-16	USA	Apollo 16	Crewed Landing	Moon
1972-12-07	USA	Apollo 17	Crewed Landing	Moon
1973-01-08	USSR	Luna 21	Rover	Moon
1973-06-10	USA	Explorer 49 (RAE-B)	Orbiter	Moon
1973-07-21	USSR	Mars 4	Mars Flyby (Attempted Mars Orbiter)	Mars
1973-07-25	USSR	Mars 5	Mars Orbiter	Mars
1973-08-05	USSR	Mars 6	Mars Lander (Contact Lost)	Mars
1973-08-09	USSR	Mars 7	Mars Flyby (Attempted Mars Lander)	Mars
1973-11-04	USA	Mariner 10	Venus/Mercury Flybys	Venus
1974-06-02	USSR	Luna 22	Orbiter	Moon
1974-10-28	USSR	Luna 23	Lander	Moon
1975-06-08	USSR	Venera 9	Venus Orbiter and Lander	Venus
1975-06-14	USSR	Venera 10	Venus Orbiter and Lander	Venus
1975-08-20	USA	Viking 1	Mars Orbiter and Lander	Mars
1975-09-09	USA	Viking 2	Mars Orbiter and Lander	Mars
1976-08-14	USSR	Luna 24	Sample Return	Moon
1978-05-20	USA	Pioneer Venus 1	Venus Orbiter	Venus
1978-08-08	USA	Pioneer Venus 2	Venus Probes	Venus
1978-09-09	USSR	Venera 11	Venus Flyby Bus and Lander	Venus
1978-09-14	USSR	Venera 12	Venus Flyby Bus and Lander	Venus
1981-10-30	USSR	Venera 13	Venus Flyby Bus and Lander	Venus
1981-11-04	USSR	Venera 14	Venus Flyby Bus and Lander	Venus

(continued)

Table A3 (continued)

Launch date	Nation	Mission name	Notes	Target
1983-06-02	USSR	Venera 15	Venus Orbiter	Venus
1983-06-07	USSR	Venera 16	Venus Orbiter	Venus
1984-12-15	USSR	Vega 1	Venus Lander and Balloon/Comet Halley Flyby	Venus
1984-12-21	USSR	Vega 2	Venus Lander and Balloon/Comet Halley Flyby	Venus
1985-07-02	EUR	Giotto	Halley comet	Flyby
1988-07-07	USSR	Phobos 1	Attempted Mars Orbiter/Phobos Landers	Mars
1988-07-12	USSR	Phobos 2	Mars Orbiter/Attempted Phobos Landers	Mars
1989-05-04	USA	Magellan	Venus Orbiter	Venus
1989-10-18	USA	Galileo	Jupiter Orbiter/Probe (Venus Flyby)	Venus
1990-01-24	JPN	Hiten	Flyby and Orbiter	Moon
1992-09-25	USA	Mars Observer	Attempted Mars Orbiter (Contact Lost)	Mars
1994-01-25	USA	Clementine	Orbiter	Moon
1996-11-07	USA	Mars Global Surveyor	Mars Orbiter	Mars
1996-11-16	RUS	Mars 96	Attempted Mars Orbiter/Landers	Mars
1996-12-04	USA	Mars Pathfinder	Mars Lander and Rover	Mars
1997-10-15	USA/EUR	Cassini-Huygens	Saturn Orbiter (Venus Flyby)	Venus
1997-12-24	USA	AsiaSat 3/HGS-1	Lunar Flyby	Moon
1998-01-07	USA	Lunar Prospector	Orbiter	Moon
1998-07-03	JPN	Nozomi (Planet-B)	Mars Orbiter	Mars
1998-12-11	USA	Mars Climate Orbiter	Attempted Mars Orbiter	Mars
1999-01-03	USA	Mars Polar Lander	Attempted Mars Lander	Mars
1999-01-03	USA	Deep Space 2 (DS2)	Attempted Mars Penetrators	Mars
2001-04-07	USA	2001 Mars Odyssey	Mars Orbiter	Mars
2003-06-02	EUR	Mars Express	Mars Orbiter and Lander	Mars
2003-06-10	USA	Spirit (MER-A)	Mars Rover	Mars
2003-07-08	USA	Opportunity (MER-B)	Mars Rover	Mars
2003-09-27	EUR	SMART 1	Lunar Orbiter	Moon
2004-03-02	EUR	Rosetta	Comet Orbiter	comet 67P
2004-08-03	USA	MESSENGER	Mercury Orbiter (Two Venus Flybys)	Venus

(continued)

Table A3 (continued)

Launch date	Nation	Mission name	Notes	Target
2005-08-12	USA	Mars Reconnaisance Orbiter	Mars Orbiter	Mars
2005-11-09	EUR	Venus Express	ESA Venus Orbiter	Venus
2006-01-20	USA	New Horizons	Pluto and Kuiper Belt	Flyby
2007-08-04	USA	Phoenix	Mars Scout Lander	Mars
2007-09-14	JPN	Kaguya (SELENE)	Lunar Orbiter	Moon
2007-10-24	CHN	Chang'e 1	Lunar Orbiter	Moon
2008-10-22	IND	Chandrayaan-1	Lunar Orbiter	Moon
2009-06-17	USA	Lunar Reconnaissance Orbiter	Lunar Orbiter	Moon
2009-06-17	USA	LCROSS	Lunar Orbiter and Impactor	Moon
2010-05-20	JPN	Akatsuki	Attempted ISAS Venus Orbiter	Venus
2010-10-01	CHN	Chang'e 2	Lunar Orbiter	Moon
2011-09-10	USA	Gravity Recovery And Interior Laboratory (GRAIL)	Lunar Orbiter	Moon
2011-11-08	CHN	Yinghuo-1	Attempted Mars Orbiter	Mars
2011-11-08	RUS	Phobos-Grunt	Attempted Martian Moon Phobos Lander	Mars
2011-11-26	USA	Mars Science Laboratory	Mars Rover	Mars
2013-09-06	USA	Lunar Atmosphere and Dust Environment Explorer	Lunar Orbiter	Moon
2013-11-05	IND	Mangalyaan	ISRO (India) Mars Orbiter	Mars
2013-11-18	USA	MAVEN	Mars Scout Mission Orbiter	Mars
2013-12-01	CHN	Chang'e 3	Lunar Lander and Rover	Moon
2016-03-14	EUR	ExoMars TGO	Mars Orbiter and Lander	Mars
2016-09-08	USA	OSIRIS-REx	Asteroid orbiter and lander	Asteroid 101955 Bennu

Table A4 Processing levels of planetary data: the definition of processing levels might be slightly confusing

Description	NASA	CODMAC	PDS3	PDS4	Isis3
Received telemetry data	Packet data	L1 (raw)	–	Telemetry	–
Reconstructed, unprocessed data	L0	L2	–	Raw	–
L0 with ancillary information	L1A	L3	EDR	Partially processed	L0
L1A processed to sensor units	L1B	L4	CDR	Partially processed	L1
Derived physical units per each sensor unit (pixel) of L1B	L2	L5 (derived)	RDR	Calibrated	L2
Variables mapped at uniform time and space scales	L3	L5	DDR	Derived	L3+
Model outputs or derived data	L4	L5	DDR	Derived	L3+

NASA terminology is similar to that currently used for Earth Observation Remote Sensing data. Terminology in Fig. A1 corresponds to the last column of this table

Data Sources

Spacecraft data used in planetary geological studies are available free of charge on the public domain, after a variable embargo period—in general of few months—where experiment teams have exclusive access to data. Software and extensive documentation are typically distributed along with data, but tools and their availability are very variable across experiments and missions. This section contains some pointers to data and documentation. The level of long-term support varies: archives are long-term preserved as well as institutionally supported tools, which are also long-term supported. The software tool scenario is rapidly changing, though.

Planetary Data Archives

Planetary Data Systems and (sample) analogue archives worldwide typically offer long term storage, curation and availability of data or samples returned by spacecrafts. The amount of extraterrestrial samples is limited, but data are steadily growing and from the few Gigabytes of total digital data holdings of few decades ago are moving towards Petabytes, several order of magnitude more. Main space agencies maintain archives from data returned by their respective missions.

NASA Planetary Data System Nodes

The NASA Planetary Data System offers data and related documentation and tools typically by broad disciplines and areas or experiment types, i.e. through several

nodes (e.g. Geosciences, Imaging Atmospheres, Small Bodies, Rings). The starting point to access all PDS resources is the NASA PDS main page.[3]

Data can be available on more than one node and search functions are available in all of them as well as from centralised (mainly web) interfaces. The most relevant node for geological analyses is the PDS Geosciences Node hosted at Washington University in St. Louis.[4] Originally data were distributed to scientists on physical archives (first CD, then DVD-based), lately all data are distributed online-only, although the term *volume* is still used.

Rover-based data are geometrically much more complex and search of data by time of observation and activity along a path/traverse is usually easier to explore and use those data. The PDS Geosciences NODE Analyst's notebook[5] provides access to Apollo, MER and MSL and more.

ESA Planetary Science Archive

ESA hosts all data coming from its planetary missions on the PSA,[6] following PDS standards. Most data from ESA PSA are also mirrored on PDS nodes (e.g. MEX HRSC). PSA is hosted in a single location at the ESAC establishment of ESA (together with astronomy data archives).

Processing Levels

Data acquired by spacecraft are returned most of the times not as *science-ready* products. The nomenclature of the different levels of calibration and processing can slightly vary, but its relative order does not (e.g. from PDS), therefore a higher level number corresponds to more science-usable data or, *higher-level data products* (Table A4).

See Figs. A1 and A2.

Web Services

The number of web services providing access, visualisation and analytics for planetary data is constantly growing, and the individual services fast evolving. Data search and discovery, possible from PDS and PDS, is enhanced within the

[3] https://pds.nasa.gov.

[4] http://pds-geosciences.wustl.edu.

[5] http://an.rsl.wustl.edu.

[6] http://www.cosmos.esa.int/web/psa.

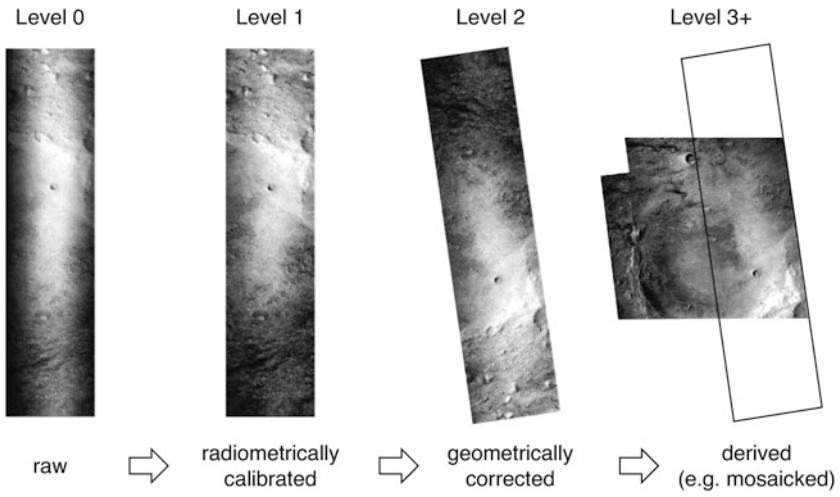

Fig. A1 Processing level examples for an imaging experiment. Levels indicated are using USGS Isis3 naming conventions. For comparison see Table A4

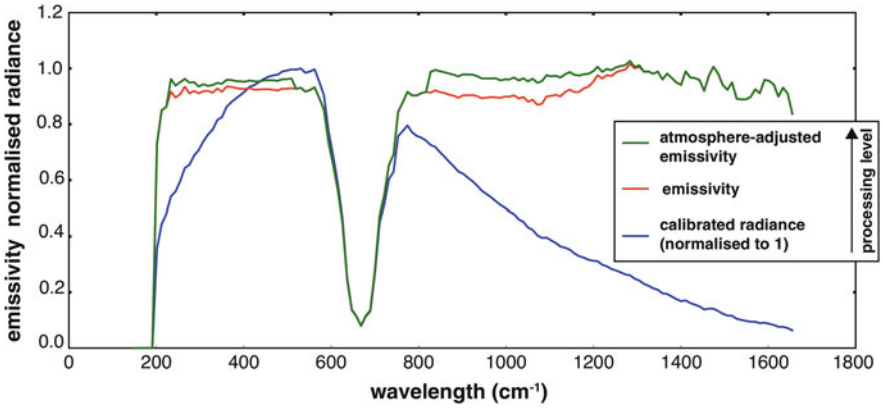

Fig. A2 Processing level examples for a spectrometer (NASA MGS TES), in this case, non-imaging: TES spectra undergo various level of processing, after being corrected for instrumental, systematic and atmospheric effects (source: MGS TES, M. D'Amore)

Planetary Virtual Observatory (VO) of EuroPlanet VESPA[7]: the VO approach, originally developed for Astronomy, allows powerful data search capabilities across an arbitrary number of archives.

[7] http://europlanet-vespa.eu.

Web mapping services (such as Web GIS systems) are widespread. Several are provided by USGS Astrogeology.[8]

Tool Directories

Several directories for planetary data analysis tools are maintained, e.g. by the PDS Geosciences Node,[9] and, of a more general nature on PDS[10] and IPDA.[11]

Imaging Tools

Video Image Communication and Retrieval (VICAR)

VICAR, originally developed since the 1960s at JPL and used for processing data from several missions, has been recently open-sourced. Its architecture influenced several later processing system such as USGS ISIS. Several pipelines use VICAR or its customisations adaptation for delivering higher-level data products to archives, e.g. MEX HRSC.

VICAR is available from JPL.[12]

USGS Integrated Software for Imaging and Spectrometers (ISIS)

ISIS[13] is a modular system developed by the USGS Astrogeology Branch and it supports several experiments on board NASA missions and beyond (e.g. ESA, ISRO, JAXA). It consists of a large set of programs to import, handle, calibrate radiometrically and geometrically planetary data from imaging cameras and spectrometers. Extensive documentation and user support is provided by USGS.

Processing of data with ISIS starts from EDR data, i.e. neither radiometrically nor geometrically calibrated data. Radiometric and geometric calibration are performed in sequence. Once imagery is map-projected it can, for example be mosaicked or further processed. The workflow to produce digital image maps is simplified in Fig. A1.

In most cases, metadata either needed or produced by the processing chains are contained in separate labels, with standards that are out of scope for this appendix,

[8] http://astrowebmaps.wr.usgs.gov/webmapatlas/Layers/maps.html.

[9] http://pds-geosciences.wustl.edu/tools/.

[10] https://pds.jpl.nasa.gov/tools/.

[11] https://planetarydata.org/services/registry.

[12] http://www-mipl.jpl.nasa.gov/vicar_open.html.

[13] https://isis.astrogeology.usgs.gov.

but contained on the PDS standard documents.[14] In the simplest case, they appear as *keyword* = *value*.

The PDS label of a sample imaging experiment (MRO CTX in this case) is something like:

```
PDS_VERSION_ID = PDS3
FILE_NAME = "D21_035563_1987_XN_18N282W.IMG"
RECORD_TYPE = FIXED_LENGTH
RECORD_BYTES = 5056
FILE_RECORDS = 20481
LABEL_RECORDS = 1
^IMAGE = 2
SPACECRAFT_NAME = MARS_RECONNAISSANCE_ORBITER
INSTRUMENT_NAME = "CONTEXT CAMERA"
INSTRUMENT_HOST_NAME = "MARS RECONNAISSANCE ORBITER"
MISSION_PHASE_NAME = "ESP"
TARGET_NAME = MARS
INSTRUMENT_ID = CTX
PRODUCER_ID = MRO_CTX_TEAM
DATA_SET_ID = "MRO-M-CTX-2-EDR-L0-V1.0"
PRODUCT_CREATION_TIME = 2014-07-01T20:11:28
SOFTWARE_NAME = "makepds05 $Revision: 1.16 $"
UPLOAD_ID = "UNK"
ORIGINAL_PRODUCT_ID = "4A_04_10B2034C00"
PRODUCT_ID = "D21_035563_1987_XN_18N282W"
START_TIME = 2014-02-26T14:46:46.527
STOP_TIME = 2014-02-26T14:47:24.963
SPACECRAFT_CLOCK_START_COUNT = "1077893243:194"
SPACECRAFT_CLOCK_STOP_COUNT = "N/A"
FOCAL_PLANE_TEMPERATURE = 293.3 <K>
SAMPLE_BIT_MODE_ID = "SQROOT"
OFFSET_MODE_ID = "197/200/187"
LINE_EXPOSURE_DURATION = 1.877 <MSEC>
SAMPLING_FACTOR = 1
SAMPLE_FIRST_PIXEL = 0
RATIONALE_DESC = "Ride-along with HiRISE"
DATA_QUALITY_DESC = "OK"
ORBIT_NUMBER = 35563
OBJECT = IMAGE
LINES = 20480
LINE_SAMPLES = 5056
LINE_PREFIX_BYTES = 0
LINE_SUFFIX_BYTES = 0
SAMPLE_TYPE = UNSIGNED_INTEGER
SAMPLE_BITS = 8
SAMPLE_BIT_MASK = 2#11111111#
CHECKSUM = 16#13621F48#
END_OBJECT = IMAGE
END
```

[14] https://pds.nasa.gov/tools/standards-reference.shtml.

The PDS label contains basic image metadata, such as its size in pixels, that are used, together with ancillary data, such as SPICE geometrical information, to perform computations of various kind, e.g. map-projecting imagery.

As an example, the corresponding subset of a level-2 (radiometrically and geometrically calibrated) Isis3 cube is:

```
Object = IsisCube
  Object = Core
    StartByte   = 65537
    Format      = Tile
    TileSamples = 128
    TileLines   = 128

    Group = Dimensions
      Samples = 9278
      Lines   = 24954
      Bands   = 1
    End_Group

    Group = Pixels
      Type       = Real
      ByteOrder  = Lsb
      Base       = 0.0
      Multiplier = 1.0
    End_Group
  End_Object

  Group = Instrument
    SpacecraftName        = Mars_Reconnaissance_Orbiter
    InstrumentId          = CTX
    TargetName            = Mars
    MissionPhaseName      = ESP
    StartTime             = 2014-02-26T14:46:46.527
    SpacecraftClockCount  = 1077893243:194
    OffsetModeId          = 197/200/187
    LineExposureDuration  = 1.877 <MSEC>
    FocalPlaneTemperature = 293.3 <K>
    SampleBitModeId       = SQROOT
    SpatialSumming        = 1
    SampleFirstPixel      = 0
  End_Group

  Group = Archive
    DataSetId           = MRO-M-CTX-2-EDR-L0-V1.0
    ProductId           = D21_035563_1987_XN_18N282W
    ProducerId          = MRO_CTX_TEAM
    ProductCreationTime = 2014-07-01T20:11:28
    OrbitNumber         = 35563
  End_Group
```

```
  Group = BandBin
    FilterName = BroadBand
    Center     = 0.65 <micrometers>
    Width      = 0.15 <micrometers>
  End_Group

  Group = Kernels
    NaifFrameCode              = -74021
    LeapSecond                 = lsk/naif0011.tls
    TargetAttitudeShape        = pck/pck00009.tpc
    TargetPosition             = (Table, spk/de405.bsp)
    InstrumentPointing         = (Table,
                                  ck/mro_sc_psp_140225_140303.bc,
                                  fk/mro_v15.tf)
    Instrument                 = Null
    SpacecraftClock            = ..sclk/MRO_SCLKSCET....tsc
    InstrumentPosition         = (Table, pk/mro_psp30.bsp)
    InstrumentAddendum         = iak/mroctxAddendum005.ti
    ShapeModel                 = dems/molaMars.cub
    InstrumentPositionQuality = Reconstructed
    InstrumentPointingQuality = Reconstructed
    CameraVersion              = 1
  End_Group

  Group = Radiometry
    FlatFile = calibration/ctxFlat_0002.cub
    iof      = 2.07298495391369e-04
  End_Group

  Group = Mapping
    ProjectionName     = EQUIRECTANGULAR
    CenterLongitude    = 0.0
    TargetName         = Mars
    EquatorialRadius   = 3396190.0 <meters>
    PolarRadius        = 3396190.0 <meters>
    LatitudeType       = Planetocentric
    LongitudeDirection = PositiveEast
    LongitudeDomain    = 360
    MinimumLatitude    = 17.529197173545
    MaximumLatitude    = 19.634177543578
    MinimumLongitude   = 77.520831823657
    MaximumLongitude   = 78.303405709876
    UpperLeftCornerX   = 4595020.0 <meters>
    UpperLeftCornerY   = 1163810.0 <meters>
    PixelResolution    = 5.0 <meters/pixel>
    Scale              = 11854.939504661 <pixels/degree>
    CenterLatitude     = 0.0
  End_Group

[...]
```

The substantially increased amount of items is generated by the processing pipeline and it reflects the increased information content of the data, including

also geometrical (mapping) information as well as a record of used ancillary data (kernels), useful for reconstructing and reproducing the processing steps applied to the data product.

NASA Ames Stereo Pipeline (ASP)

Stereogrammetry from remote sensing imagery can produced digital terrain models usable as topographic base for detailed geological observations (large scale topography is often guaranteed by global low-resolution laser altimetry, on most terrestrial planets at least).

The NASA Stereo Pipeline,[15] developed at the AMES research center is an actively developed, powerful Open-Source photogrammetric package, capable of working with data from most planetary missions. It is very actively supported via a discussion group/mailing list.

JMars/JMoon, etc.

JMars[16] and its companion tools for other planetary bodies (e.g. Moon) is a popular GIS system. It provided access to a large base of remote sensing data and it is actively developed.

Tools Usable on the Web

The tools above, especially ISIS, can be also run as web services on demand, producing e.g. mosaics, resampling or reprojecting data.

Processing on the Web (POW)[17] allows such functionalities and data processing and delivery on demand (not in real time). On demand processing is also provided by the e-Mars MarSI system.[18] A recent development of on-line real-time analytics on planetary data is constituted by PlanetServer/EarthServer,[19] using OGC WCPS to query data. Access and download of data such as MEX HRSC can be achieved using tools developed by experiment teams and freely accessible, such as HRSC Maps orbit locator.[20]

There are numerous other tools and services available from USGS and others. This number is very likely to increase on very short time scales. Please refer to the

[15] http://ti.arc.nasa.gov/tech/asr/intelligent-robotics/ngt/stereo/.

[16] https://jmars.mars.asu.edu.

[17] http://astrocloud.wr.usgs.gov.

[18] https://emars.univ-lyon1.fr/MarsSI/.

[19] http://planetserver.eu.

[20] http://maps.planet.fu-berlin.de.

USGS Astrogeology Branch Mercator Lab[21] as well as the GitHub repository of the present book.[22]

Additionally, *citizen science* projects and tools exist, either based on surface change detection using multitemporal data or on texture analysis, such as Planet4,[23] supported by the HiRISE experiment team.

Documentation and Resources

Documentation on planetary exploration is available on a variety of long-term maintained web sites (a subset is reported below), as well as on more volatile media (not reported here):

NASA Space Science Coordinated Data Archive (NSSDCA[24]) contains up-to-date information on planetary missions and facts, also beyond NASA.

Tutorials and Workshops

Planetary data workshops (also known as data user workshops, or alike) are available from PDS, PSA and additional parties. They tend to be updated periodically.

An up to date list of workshop relevant for Planetary Geology is maintained on the PDS Geosciences Node.[25] ESA PSA has a dedicated workshop page.[26] Planetary data analysis and mapping workshops are also regularly run at USGS and materials collected for anyone to use.[27]

Few recent representative workshops are listed below (the list is not exhaustive):

- MEX HRSC / OMEGA data workshop (2007, 2008)
- MEX MARSIS (2008)
- MRO CRISM data user workshop (2009, 2012)
- Chandrayaan M3 data workshop (2010)
- MRO SHARAD data workshop (2014)
- USGS Planetary data workshop (2012, 2015)
- ESAC Planetary GIS data workshop (2015)
- MSL ChemCam (2015)

[21] http://astrogeology.usgs.gov/facilities/mrctr-gis-lab.

[22] https://github.com/openplanetary/planetarygeology-book.

[23] https://www.planetfour.org.

[24] http://nssdc.gsfc.nasa.gov/planetary.

[25] http://pds-geosciences.wustl.edu/workshops/.

[26] http://www.sciops.esa.int/index.php?project=PSA&page=workshops.

[27] http://astrogeology.usgs.gov/groups/Planetary-Data-Workshop.

Additional Resources

Additional resources include discussion groups or online communities such as OpenPlanetary,[28] Isis Support,[29] NASA Stereo Pipeline mailing list.[30]

[28] https://github.com/openplanetary.

[29] http://isis.astrogeology.usgs.gov/fixit.

[30] stereo-pipeline@lists.nasa.gov.

Locations

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8

Persons

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8

Subjects

A.P. Rossi, S. van Gasselt (eds.), *Planetary Geology*, Springer Praxis Books,
DOI 10.1007/978-3-319-65179-8