6

Impact cratering

> The dominant surface features of the Moon are approximately circular depressions, which may be designated by the general term craters … Solution of the origin of the lunar craters is fundamental to the unraveling of the history of the Moon and may shed much light on the history of the terrestrial planets as well.
>
> E. M. Shoemaker (1962)

Impact craters are the dominant landform on the surface of the Moon, Mercury, and many satellites of the giant planets in the outer Solar System. The southern hemisphere of Mars is heavily affected by impact cratering. From a planetary perspective, the *rarity or absence* of impact craters on a planet's surface is the exceptional state, one that needs further explanation, such as on the Earth, Io, or Europa. The process of impact cratering has touched every aspect of planetary evolution, from planetary accretion out of dust or planetesimals, to the course of biological evolution.

The importance of impact cratering has been recognized only recently. E. M. Shoemaker (1928–1997), a geologist, was one of the first to recognize the importance of this process and a major contributor to its elucidation. A few older geologists still resist the notion that important changes in the Earth's structure and history are the consequences of extraterrestrial impact events. The decades of lunar and planetary exploration since 1970 have, however, brought a new perspective into view, one in which it is clear that high-velocity impacts have, at one time or another, affected nearly every atom that is part of our planetary system. Impact cratering is crucially important for the accumulation of the planets in the first place and has played major roles from the formation of the most ancient planetary landscapes to the creation and maintenance of the modern regolith of airless bodies. In an important sense, impact cratering is the most fundamental geologic process in the Solar System.

6.1 History of impact crater studies

Craters were discovered in 1610 when Galileo pointed his first crude telescope at the Moon. Galileo recognized the raised rims and central peaks of these features, but described them only as circular "spots" on the Moon. Although Galileo himself did not record an opinion on how they formed, astronomers argued about their origin for the next three centuries.

Astronomer J. H. Schröter first used the word "crater" in a non-genetic sense in 1791. Until the 1930s most astronomers believed the Moon's craters were giant extinct volcanoes: the impact hypothesis, proposed sporadically over the centuries, did not gain a foothold until improving knowledge of impact physics showed that even a moderately oblique high-speed impact produces a circular crater, consistent with the observed circularity of nearly all of the Moon's craters. Even so, many astronomers clung to the volcanic theory until high-resolution imagery and *in situ* investigation of the Apollo program in the early 1970s firmly settled the issue in favor of an impact origin for nearly every lunar crater. In the current era spacecraft have initiated the remote study of impact craters on other planets, beginning with Mariner 4's unexpected discovery of craters on Mars on July 15, 1965. Since then craters have been found on almost every other solid body in the Solar System.

Meteor Crater, Arizona, was the first terrestrial structure shown unambiguously to be of impact origin. D. M. Barringer (1860–1929) investigated this 1 km diameter crater and its associated meteoritic iron in detail from 1906 until his death in 1929. After Barringer's work a large number of small impact structures resembling Meteor Crater have been found. Impact structures larger than about 5 km in diameter were first described as "cryptovolcanic" because they showed signs of violent upheaval but were not associated with the eruption of volcanic materials. J. D. Boon and C. C. Albritton in 1937 proposed that these structures were really caused by impacts, although final proof had to wait until the 1960s when the shock-metamorphic minerals coesite and stishovite proved that the Ries Kessel in Germany and subsequently many other cryptovolcanic structures are the result of large meteor impacts.

Finally, theoretical and experimental work on the mechanics of cratering began during World War II and was extensively developed in later years. This work was spurred partly by the need to understand the craters produced by nuclear weapons and partly by the fear that the "meteoroid hazard" to space vehicles would be a major barrier to space exploration. Computer studies of impact craters were begun in the early 1960s. A vigorous and highly successful experimental program to study the physics of impact was initiated by D. E. Gault (1923–1999) at NASA's Ames facility in 1965.

These three traditional areas of astronomical crater studies, geological investigation of terrestrial craters, and the physics of cratering have blended together in the post-Apollo era. Traditional boundaries have become blurred as extraterrestrial craters are subjected to direct geologic investigation, the Earth's surface is scanned for craters using satellite images, and increasingly powerful computers are used to simulate the formation of both terrestrial and planetary craters on all size scales. The recent proposals that the Moon was created by the impact of a Mars-sized protoplanet with the proto-Earth 4.5 Gyr ago and that the Cretaceous era was ended by the impact of a 15 km diameter asteroid or comet indicate that the study of impact craters is far from exhausted and that new results may be expected in the future.

6.2 Impact crater morphology

Fresh impact craters can be grossly characterized as "circular rimmed depressions." Although this description can be applied to all craters, independent of size, the detailed

form of craters varies with size, substrate material, planet, and age. Craters have been observed over a range of sizes varying from 0.1 μm (microcraters first observed on lunar rocks brought back by the Apollo astronauts) to the more than 2000 km diameter Hellas Basin on Mars. Within this range a common progression of morphologic features with increasing size has been established, although exceptions and special cases are common.

6.2.1 Simple craters

The classic type of crater is the elegant bowl-shaped form known as a "simple crater" (Figure 6.1a). This type of crater is common at sizes less than about 15 km diameter on the Moon and 3 to about 6 km on the Earth, depending on the substrate rock type. The interior of a simple crater has a smoothly sloping parabolic profile and its rim-to-floor depth is about 1/5 of its rim-to-rim diameter. The sharp-crested rim stands about 4% of the crater diameter above the surrounding plain, which is blanketed with a mixture of ejecta and debris scoured from the pre-existing surface for a distance of about one crater diameter from the rim. The thickness of the ejecta falls off as roughly the inverse cube of distance from the rim. The surface of the ejecta blanket is characteristically hummocky, with mounds and hollows alternating in no discernible pattern. Fields of small secondary craters and bright rays of highly pulverized ejecta that extend many crater diameters away from the primary may surround particularly fresh simple craters. Meteor Crater, Arizona, is a slightly eroded representative of this class of relatively small craters. The floor of simple craters is underlain by a lens of broken rock, "breccia," which slid down the inner walls of the crater shortly following excavation. This breccia typically includes representatives from all the formations intersected by the crater and may contain horizons of melted or highly shocked rock. The thickness of this breccia lens is typically 1/2 to 1/3 of the rim-to-floor depth.

6.2.2 Complex craters

Lunar craters larger than about 20 km diameter and terrestrial craters larger than about 3 km have terraced walls, central peaks, and at larger sizes may have flat interior floors or internal rings instead of central peaks. These craters are believed to have formed by the collapse of an initially bowl-shaped "transient crater," and because of this these more complicated structures are known as "complex craters" (Figure 6.1b). The transition between simple and complex craters has now been observed on the Moon, Mars, Mercury, and the Earth, as well as on some of the icy satellites in the outer Solar System. In general the transition diameter scales as g^{-1}, where g is the acceleration of gravity at the planet's surface, although the constant in the scaling rule is not the same for icy and rocky bodies. This is consistent with the idea that complex craters form by collapse, with icy bodies having only about 1/3 the strength of rocky ones. The floors of complex craters are covered by melted and highly shocked debris, and melt pools are sometimes seen in depressions in the surrounding ejecta blanket. The surfaces of the terrace blocks tilt outward into the crater walls, and melt pools are also common in the depressions thus formed. The most notable

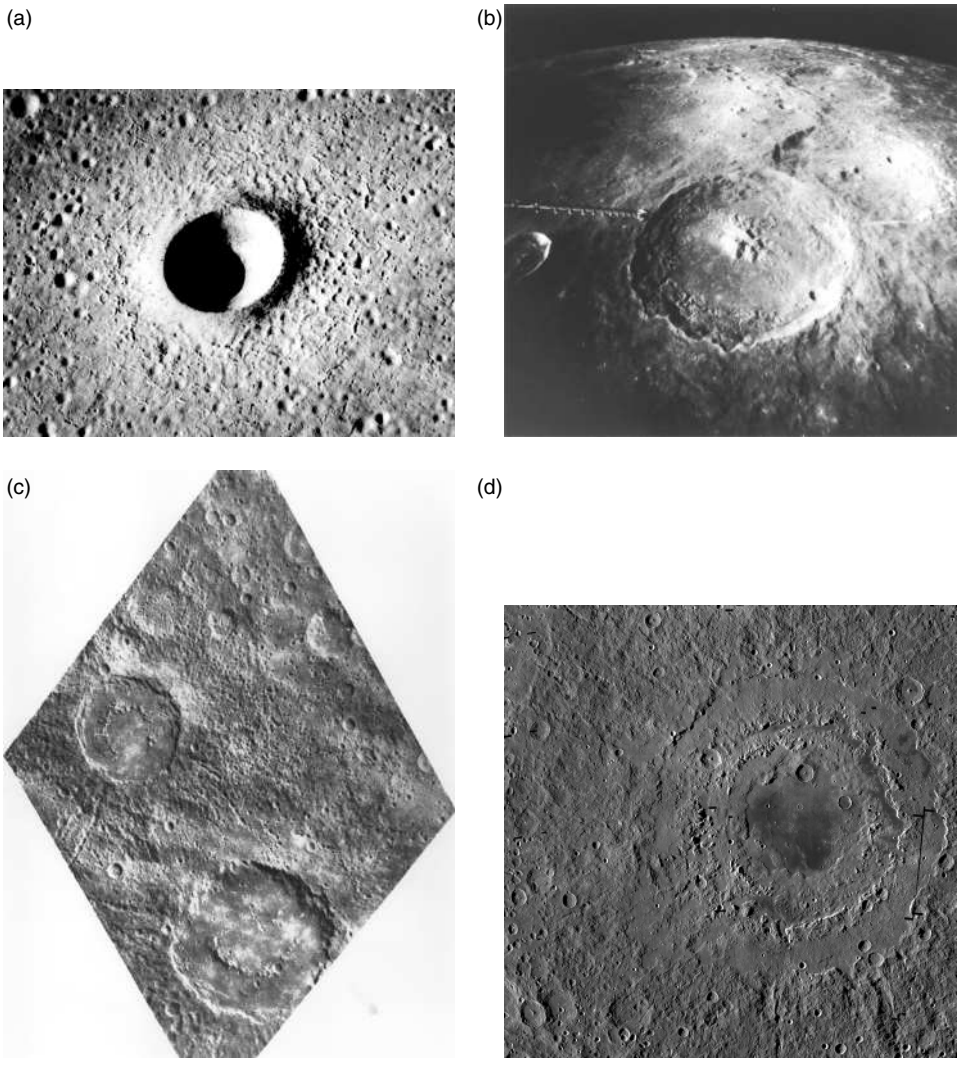

Figure 6.1 Impact crater morphology as a function of increasing size. (a) Simple crater: 2.5 km diameter crater Linné on the Moon (Apollo 15 Panometric Photo strip 9353). (b) Complex crater with central peak: 102 km diameter crater Theophilus on the Moon (Apollo 16 Hasselblad photo 0692). (c) Complex crater with internal ring: Mercurian craters Strindberg (165 km diameter) to the lower right and Ahmad Baba (115 km) to the upper left (Mariner 10 FDS 150, rectified). (d) Multiring basin: 620 km diameter (of most prominent ring) Orientale basin on the Moon (LROC WAC mosaic. Full width of image PIA13225 is 1350 km. NASA/GSFC/ASU).

structural feature of complex craters is the uplift beneath their centers. The central peaks contain material that is pushed upward from the deepest levels excavated by the crater. Study of terrestrial craters has shown that the amount of structural uplift h_{su} is related to the final crater diameter D by:

$$h_{su} = 0.06\, D^{1.1} \tag{6.1}$$

where all distances are in kilometers. The diameter of the central peak complex is roughly 22% of the final rim-to-rim crater diameter in craters on all the terrestrial planets.

Complex craters are shallower than simple craters of equal size and their depth increases slowly with increasing crater diameter. On the Moon, the depth of complex craters increases from about 3 km to only 6 km while crater diameters range from 20 to 400 km. Rim height also increases rather slowly with increasing diameter because much of the original rim slides into the crater bowl as the wall collapses. Complex craters are thus considerably larger than the transient crater from which they form: estimates suggest that the crater diameter may increase as much as 60% during collapse. A useful scaling relation suggests that the rim-to-rim diameter of a complex crater is related to the transient crater D_t diameter by:

$$D = 1.17 \frac{D_t^{1.13}}{D_{s-c}^{0.13}} \tag{6.2}$$

where $D_{s\text{-}c}$ is the diameter at the simple to complex transition, about 3.2 km on the Earth and 15 km on the Moon.

As crater size increases, the central peaks characteristic of smaller complex craters give way to a ring of mountains (Figure 6.1c). This transition takes place at about 140 km diameter on the Moon and about 20 km diameter on the Earth, again following a g^{-1} rule. Known as "peak-ring craters," the central ring is generally about 0.5 of the rim-to-rim diameter of the crater on all the terrestrial planets.

The ejecta blankets of complex craters are generally similar to those of simple craters, although radial troughs and ridges replace the "hummocky" texture characteristic of simple craters as size increases. Fresh complex craters also have well-developed fields of secondary craters, including frequent clusters and "herringbone" chains of closely associated, irregular, secondary craters. Very fresh craters, such as Copernicus and Tycho on the Moon, have far-flung bright ray systems.

6.2.3 Multiring basins

The very largest impact structures are characterized by multiple concentric circular scarps, and are, hence, known as "multiring basins." The most famous such structure is the 930 km diameter Orientale basin on the Moon (Figure 6.1d), which has at least four nearly complete rings of inward-facing scarps. Although opinion on the origin of the rings still varies, most investigators feel that the scarps represent circular faults that slipped shortly after the crater was excavated. There is little doubt that multiring basins are caused by impacts: most of them have recognizable ejecta blankets characterized by a radial ridge-and-trough pattern. The ring diameter ratios are often tantalizingly close to multiples of $\sqrt{2}$, although no one has yet suggested a convincing reason for this relationship.

In contrast to the simple/complex and central peak/internal ring transitions discussed above, the transition from complex craters to multiring basins is not a simple function of

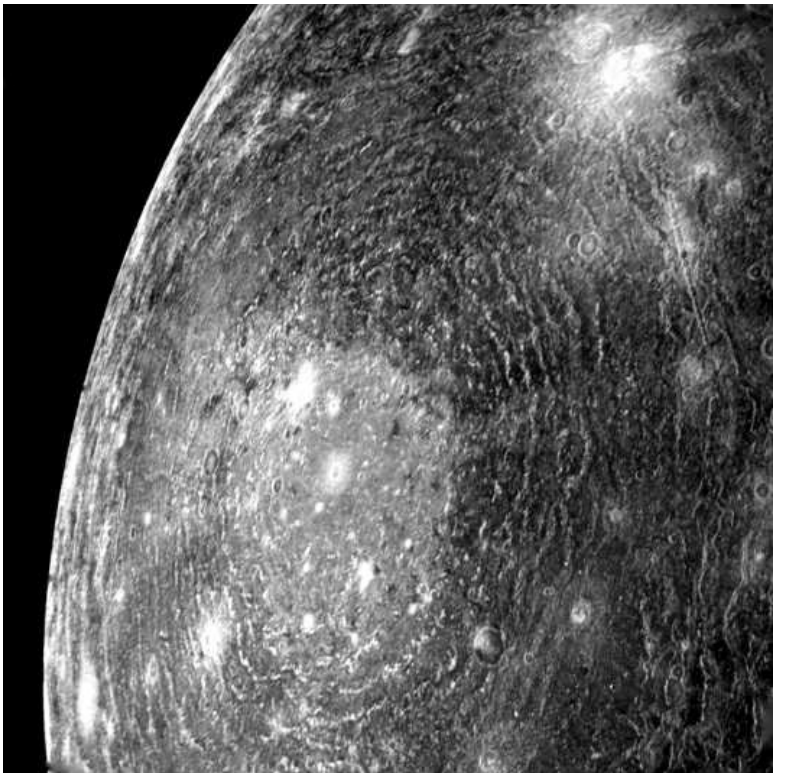

Figure 6.2 The Valhalla basin on Callisto. The original impact was within the central bright patch, which is 300 km in diameter and may represent ejecta from a still smaller (now unrecognizable) crater. This central zone is surrounded by an annulus of sinuous ridges, which in turn is surrounded by an annulus of trough-like grabens, which can be recognized up to 2000 km from the basin center. Voyager 1 mosaic PIA02277. NASA/JPL.

g^{-1}. Although multiring basins are common on the Moon, where the smallest has a diameter of 410 km, none at all has been recognized on Mercury, with its two times larger gravity, even though the largest crater, Caloris Basin, is 1540 km in diameter. The situation on Mars has been confused by erosion, but it is difficult to make a case that even the 1200 km diameter Argyre Basin is a multiring structure. A very different type of multiring basin is found on Jupiter's satellite Callisto, where the 4000 km diameter Valhalla basin (Figure 6.2) has dozens of closely spaced rings that appear to face outward from the basin center. Another satellite of Jupiter, Ganymede, has both Valhalla-type and Orientale-type multiring structures. Since gravity evidently does not play a simple role in the complex crater/multiring basin transition, some other factor, such as the internal structure of the planet, may have to be invoked to explain the occurrence of multiring basins. The formation of such basins is currently a topic of active research.

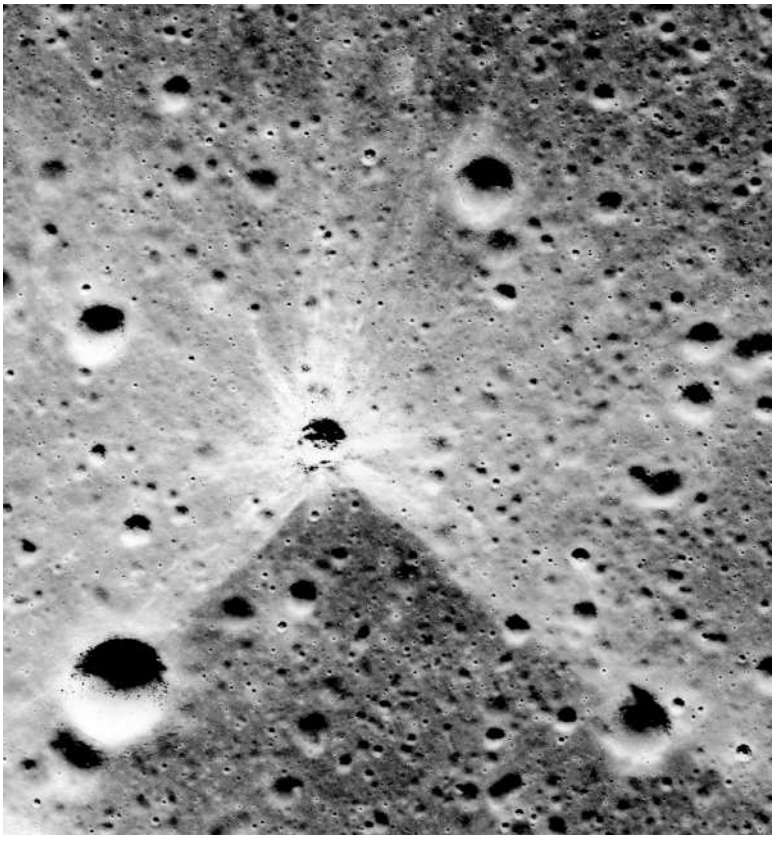

Figure 6.3 Distinctive asymmetrical ejecta surrounds a 370 m diameter crater on the lunar mare near the Linné crater 25° N, 17° W. This pattern is typical of ejecta surrounding a crater formed by an impact at an angle between 20° and 45° from the horizontal. The crater itself is still circular, but there is an uprange wedge in which very little ejecta is present. Portion of Apollo 15 Panometric image AS15-P-9337.

6.2.4 Aberrant crater types

On any planetary surface a few craters can always be found that do not fit the simple size–morphology relation described above. These are generally believed to be the result of unusual conditions of formation in either the impacting body or the planet struck. Circular craters with asymmetric ejecta blankets (Figure 6.3) or elliptical craters with "butterfly-wing" ejecta patterns are the result of very low impact angles. Although moderately oblique impacts yield circular craters, at impact angles less than about 6° from the horizontal the final crater becomes elongated in the direction of flight. Small, apparently concentric, craters or craters with central dimples or mounds on their floors are the result of impacts into a weak layer underlain by a stronger one. The ejecta blankets of some Martian craters show petal-like flow lobes that are believed to indicate the presence of liquid water in the

excavated material. Craters on Ganymede and Callisto have central pits at a diameter where internal rings would be expected on other bodies. The explanation for these pits is still unknown. In spite of these complications, however, the simple size–morphology relation described above provides a simple organizing principle into which most impact craters can be grouped.

6.2.5 Degraded crater morphology

All crater morphologies are observed in either "fresh," pristine landforms or as *erosionally degraded* forms. The dominant degradational process determines the detailed changes in crater morphology. Other impact craters are often the exclusive agent of degradation on airless bodies like the Moon (although burial by floods of lava can be locally important there as well). This form of erosion fuels a variety of surface creep (see Section 8.1 for more detail). Sharp terrain features such as crater rims are rounded and battered out of line by smaller impacts, crater bowls are gradually filled and slopes become gentler. At the extreme limit, craters may fade into invisibility as their place is occupied by large numbers of overlapping craters. Moon mappers have established "degradation classes" for lunar craters that range from fresh to nearly invisible and depend upon the initial crater size. Used in conjunction with crater density data, the numbers of craters in different degradation classes can be used to infer the age and cratering history of a given site for different populations of impactors.

On ancient Mars, fluvial processes dissected impact craters by gullying and channel formation. Old craters there were filled with sediment and lava. Wind-blown sand and dust fill small craters on Mars today and erode their rims into crenulated yardang ridges.

Erosion on Earth is so active that craters are among the most rare landforms. Fluvial deposition fills in closed depressions, such as crater bowls, and fluvial erosion gullies rims and quarries away ejecta blankets. Many of the craters that are fortunately preserved were once completely buried, preserving them, and are only now being exhumed: The Ries crater in Germany is an example of this fortunate circumstance. Differential erosion of the various rock units etches out the present morphology of the crater to create its modern landscape.

The varieties of degraded crater morphologies are as diverse as the different agencies of erosion or deposition. Recognition of degraded crater forms must, thus, take the behavior of each process into account as observers attempt to reconstruct the original structure of an impact crater.

6.3 Cratering mechanics

The impact of an object moving at many kilometers per second with the surface of a planet initiates an orderly sequence of events that eventually produces an impact crater. Although this is really a continuous process, it is convenient to break it up into distinct stages that are each dominated by different physical processes. This division clarifies the description

of the overall cratering process, but it should not be forgotten that the different stages really grade into one another and that a perfectly clean separation is not possible. The most commonly used division of the impact cratering process is into contact and compression, excavation, and modification.

6.3.1 Contact and compression

Contact and compression is the briefest of the three stages, lasting only a few times longer than the time required for the impacting object (referred to hereafter as the "projectile") to traverse its own diameter, $t_{cc} \approx L/v_i$, where t_{cc} is the duration of contact and compression, L is the projectile diameter, and v_i is the impact velocity. During this stage the projectile first contacts the planet's surface (hereafter, "target") and transfers its energy and momentum to the underlying rocks. The specific kinetic energy (energy per unit mass, $\frac{1}{2} v_i^2$) possessed by a projectile traveling at even a few kilometers per second is surprisingly large. A. C. Gifford, in 1924, first realized that the energy per unit mass of a body traveling at 3 km/s is comparable to that of TNT. Gifford proposed the "impact-explosion analogy," which draws a close parallel between a high-speed impact and an explosion. During contact and compression the projectile plunges into the target, generating strong shock waves as the material of both objects is compressed. The strength of these shock waves can be computed from the Hugoniot equations, first derived by P. H. Hugoniot in his 1887 thesis, that relate quantities in front of the shock (subscript 0) to quantities behind the shock (no subscript):

$$\begin{aligned} \rho(U-u_p) &= \rho_0 U \\ P-P_0 &= \rho_0 u_p U \\ E-E_0 &= \frac{(P+P_0)}{2}\left(\frac{1}{\rho_0}-\frac{1}{\rho}\right). \end{aligned} \tag{6.3}$$

In these equations P is pressure, ρ is density, u_p is particle velocity behind the shock (the unshocked material is assumed to be at rest), U is the shock velocity, and E is energy per unit mass. These three equations are equivalent to the conservation of mass, momentum, and energy, respectively, across the shock front. They hold for all materials, but do not provide enough information to specify the outcome of an impact by themselves. The Hugoniot equations must be supplemented by a fourth equation, the equation of state, that relates the pressure to the density and internal energy in each material, $P = P(\rho, E)$. Alternatively, a relation between shock velocity and particle velocity may be specified, $U = U(u_p)$. Since this relation is frequently linear, it often provides the most convenient equation of state in impact processes. Thus,

$$U = c + Su_p \tag{6.4}$$

where c and S are empirical constants (c is the bulk sound speed and S is a dimensionless slope). Table 6.1 lists the measured values of c and S for a variety of materials. These

Table 6.1. *Linear shock-particle velocity equation of state parameters*

Material	ρ_0 (kg/m^3)	c (km/s)	S
Aluminum	2750	5.30	1.37
Basalt	2860	2.6	1.62
Calcite (carbonate)	2670	3.80	1.42
Coconino sandstone	2000	1.5	1.43
Diabase	3000	4.48	1.19
Dry sand	1600	1.7	1.31
Granite	2630	3.68	1.24
Iron	7680	3.80	1.58
Permafrost (water saturated)	1960	2.51	1.29
Serpentinite	2800	2.73	1.76
Water (25°C)	998	2.393	1.333
Water Ice (–15°C)	915	1.317	1.526

Data from Melosh (1989).

equations can be used to compute the maximum pressure, particle velocity, shock velocity, etc. in an impact.

Planar impact approximation. A rough estimate of the parameters describing the highest pressure portion of the contact and compression stage is obtained from the planar impact approximation (sometimes called the impedance matching solution), which is valid so long as the lateral dimensions of the projectile are small compared with the distance the shock has propagated. This approximation is, thus, valid through most of the contact and compression stage. A simultaneous solution to the Hugoniot jump equations is obtained in both the target and projectile by noting that, at the interface between the two, both the particle velocity and pressure must be the same in both bodies. Unfortunately, there is no simple formula for this approximation. The simplest expression is for the particle velocity in the target, u_t (the particle velocity in the projectile is v_i-u_t by the velocity matching condition), which is the solution of a simple quadratic equation:

$$u_t = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \tag{6.5}$$

where A, B, and C are defined as:

$$\begin{aligned} A &= \rho_{0t} S_t - \rho_{0p} S_p \\ B &= \rho_{0t} c_t + \rho_{0p} c_p + 2\rho_{0p} S_p v_i. \\ C &= -\rho_{0p} v_i (c_p + S_p v_i) \end{aligned} \tag{6.6}$$

The subscripts p and t refer to the projectile and target respectively. The above equation can be used in conjunction with the Hugoniot equations and equation of state to obtain

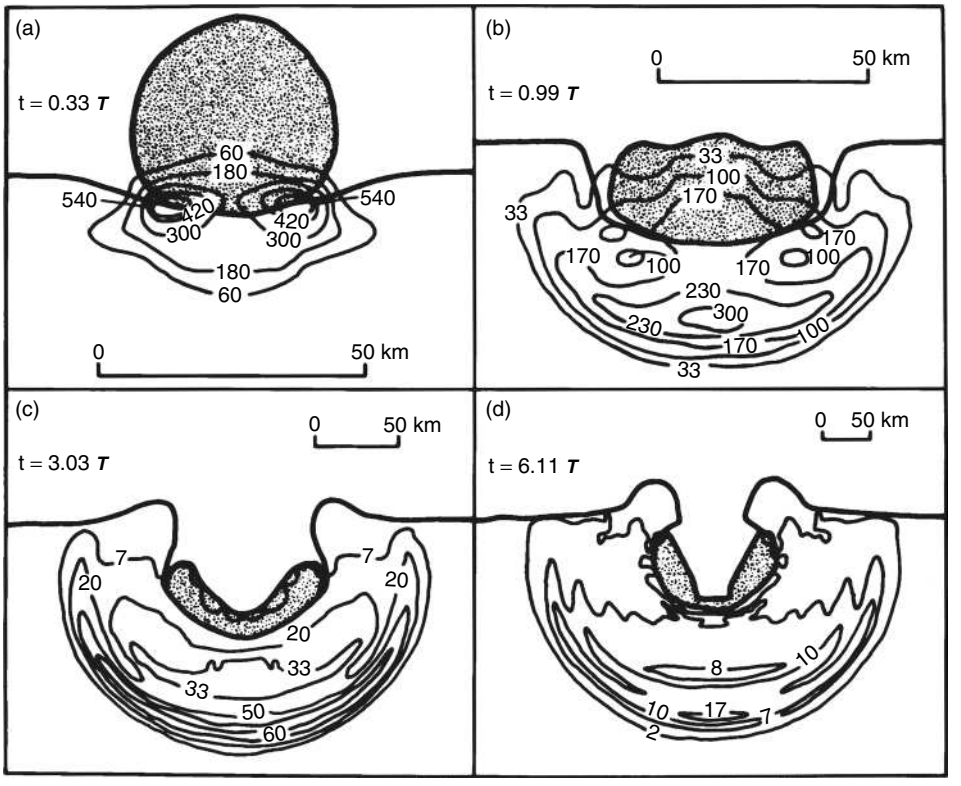

Figure 6.4 The first three frames (a to c) illustrate the evolution of shock waves in the contact and compression stages of the vertical impact of a 46.4 km diameter iron projectile on a gabbroic anorthosite target at 15 km/s. The last frame (d) is a very early phase of the excavation stage. Pressure contours are labeled in GPa, and the times given are in multiples of the time that the projectile takes to pass through its own diameter, about 3 s in this case. Note the change in length scale from frame to frame.

any other quantities of interest. Thus, the pressure behind the shock is given by $P = \rho_{ot}\, u_t\, (c_t + S_t\, u_t)$. The pressures in both the target and projectile are the same by construction of the solution.

As the projectile plunges into the target, shock waves propagate both into the projectile, compressing and slowing it, and into the target, compressing and accelerating it downward and outward (Figure 6.4). At the interface between target and projectile the material of each body moves at the same velocity. This equals 1/2 the impact velocity if they are composed of the same materials (note that in the above equation, $A = 0$ in this case, but the numerator also vanishes and the right-hand side of the equation approaches $-C/B$, which equals $v_i/2$). The shock wave in the projectile eventually reaches its back (or top) surface. When this happens, the pressure is released as the surface of the compressed projectile expands

upward and a wave of pressure relief propagates back downward toward the projectile–target interface. The contact and compression stage is considered to end when this relief wave reaches the projectile–target interface. At this time the projectile has been compressed to high pressure, often reaching hundreds of gigapascals, and upon decompression it may be in the liquid or gaseous state due to heat deposited in it during the irreversible compression process. The projectile generally carries off 50% or less of the total initial energy if the density and compressibility of the projectile and target material do not differ too much, while the balance of the energy moves into the target. This energy will eventually be expended in opening the crater as well as heating the target. The projectile–target interface at the end of contact and compression is generally less than one projectile diameter L below the original surface.

Contact and compression are accompanied by the formation of very high-velocity "jets" of highly shocked material. These jets form where strongly compressed material is close to a free surface, for example near the circle where a spherical projectile contacts a planar target. The jet velocity depends on the angle between the converging surface of the projectile and target, but may exceed the impact velocity by factors as great as 5. Jetting was initially regarded as a spectacular but not quantitatively important phenomenon in early impact experiments, where the incandescent streaks of jetted material only amounted to about 10% of the projectile's mass in vertical impacts. However, recent work on oblique impacts indicates that in this case jetting is much more important and that the entire projectile may participate in a downrange stream of debris that carries much of the original energy and momentum. Oblique impacts are still not well understood and more work needs to be done to clarify the role of jetting early in this process.

6.3.2 Excavation

During the excavation stage the shock wave created during contact and compression expands and eventually weakens into an elastic wave, while the crater itself is opened by the much slower "excavation flow." The duration of this stage is roughly given by the period of a gravity wave with wavelength equal to the crater diameter D, equal to $t_{EX} \sim (D/g)^{1/2}$, for craters whose excavation is dominated by gravity g (this includes craters larger than a few kilometers in diameter, even when excavated in hard rock). Thus, Meteor Crater, Arizona, was excavated in about 10 s, while the 1000 km diameter Imbrium Basin on the Moon took about 13 minutes to open. Shock wave expansion and crater excavation, while intimately linked, occur at very different rates and may be usefully considered separately.

The high pressure attained during contact and compression is almost uniform over a volume roughly comparable to the initial dimensions of the projectile, a region called the "isobaric core." However, as the shock wave expands away from the impact site the shock pressure declines as the initial impact energy spreads over an increasingly large volume of rock and loses energy to heating the target. The pressure P in the shock wave as a function of distance r from the impact site is given roughly by:

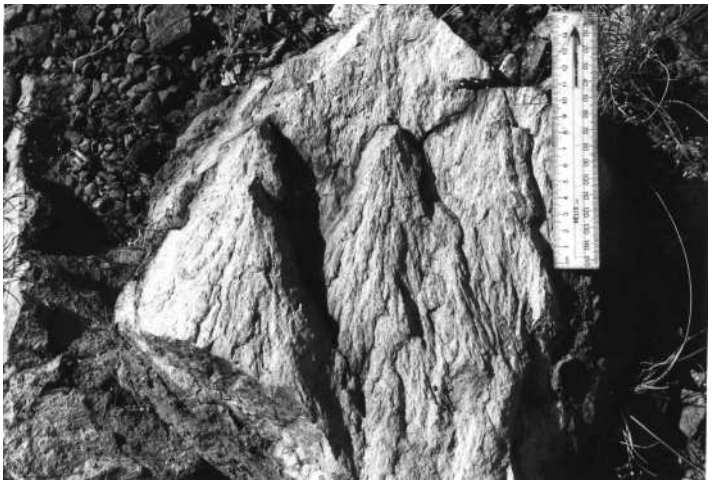

Figure 6.5 Shatter cones from the Spider Structure, Western Australia, formed in mid-Proterozoic orthoquartzite. This cone-in-cone fracture is characteristic of shattering by impact-generated shock waves. The scale bar is 15 cm long (courtesy of George Williams).

$$P = P_0\left(\frac{a}{r}\right)^n \tag{6.7}$$

where a (= $L/2$) is the radius of the projectile, P_0 is the pressure established during contact and compression, and the power n is between 2 and 4, depending on the strength of the shock wave (n is larger at higher pressures – a value n = 3 is a good general average).

Shock metamorphism. The shock wave, with a release wave immediately following, quickly attains the shape of a hemisphere expanding through the target rocks. The high-shock pressures are confined to the surface of the hemisphere: the interior has already decompressed. The shock wave moves very quickly, as fast or faster than the speed of sound, between about 6 and 10 km/s in most rocks. As rocks in the target are overrun by the shock waves, then released to low pressures, mineralogical changes take place in the component minerals. At the highest pressures the rocks may melt or even vaporize upon release. As the shock wave weakens high-pressure minerals such as coesite or stishovite arise from quartz in the target rocks, diamonds may be produced from graphite, or maskelynite from plagioclase. Somewhat lower pressures cause pervasive fracturing and "planar elements" in individual crystals. Still lower pressures create a characteristic cone-in-cone fracture called "shatter cones" (Figure 6.5),which are readily recognized in the vicinity of impact structures. Indeed, many terrestrial impact structures were first recognized from the occurrence of shatter cones. Table 6.2 lists a number of well-established shock metamorphic changes and the pressures at which they occur.

Table 6.2. *Petrographic shock indicators*

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

Data from Melosh (1989).

Spallation. The expanding shock wave encounters a special condition near the free surface. The pressure at the surface must be zero at all times. Nevertheless, a short distance below the surface the pressure is essentially equal to P, defined above. This situation results in a thin layer of surface rocks being thrown upward at very high velocity (the theoretical maximum velocity approaches the impact speed v_i). Since the surface rocks are not compressed to high pressure, this results in the ejection of a small quantity of unshocked or lightly shocked rocks at speeds that may exceed the target planet's escape velocity. Although the total quantity of material ejected by this "spall" mechanism is probably only 1–3% of the total mass excavated from the crater, it is particularly important scientifically as this is probably the origin of the recently discovered meteorites from the Moon, and of the SNC (shergottite, nakhlite, and chassignite) meteorites, which are widely believed to have been ejected from Mars.

Seismic shaking. The weakening shock wave eventually degrades into elastic waves. These elastic waves are similar in many respects to the seismic waves produced by an earthquake, although impact-generated waves contain less of the destructive shear-wave energy than earthquake waves. The seismic waves produced by a large impact may have significant effects on the target planet, creating jumbled terrains at the antipode of the impact site if they are focused by internal planetary structures, such as a low-velocity core.

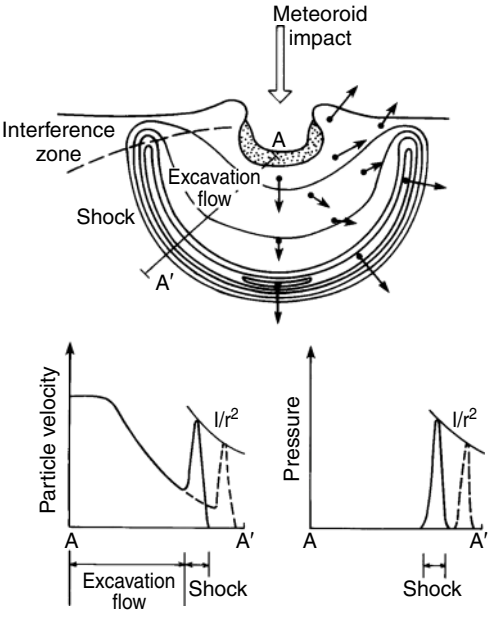

Figure 6.6 Illustration of the expanding shock wave and excavation flow following a meteorite impact. The contours in the upper part of the figure represent the pressure at some particular time after the impact. The region of high-shock pressure is isolated or "detached" on an expanding hemispherical shell. The lower graphs show profiles of particle velocity and pressure along the section AA′. The dashed lines on these graphs show the particle velocity and pressure some time later than those shown by the solid lines, and the solid curves connecting the peaks are portions of the "envelopes" of peak particle velocity and peak pressure.

This effect has been observed opposite Caloris Basinon Mercury and opposite Imbrium and Orientale on the Moon. The equivalent Richter magnitude M caused by an impact of energy E (= $1/2\ m_p v_i^2$) is given approximately by:

$$M = 0.67 \log_{10} E - 5.87. \tag{6.8}$$

Excavation mechanics. Target material engulfed by the shock wave is released a short time later. Upon release the material has a velocity that is only about 1/5 of the particle velocity in the shock wave. This "residual velocity" is due to thermodynamic irreversibility in the shock compression. It is this velocity field that eventually excavates the crater (Figure 6.6). The excavation velocity field has a characteristic downward-outward-then upward pattern that moves target material out of the crater, ejecting it at angles close to 45° at the rim. The streamlines of this flow cut across the contours of maximum shock pressure, so that material ejected at any time may contain material with a wide range of shock levels (Figure 6.7). Nevertheless, the early, fast ejecta generally contain a higher proportion of highly shocked material than the later, slower ejecta. Throughout its growth the crater is lined with highly shocked, often melted, target material.

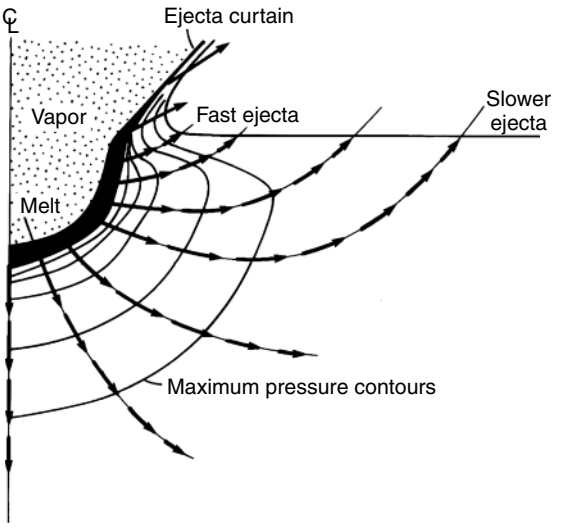

Figure 6.7 Geometry of the excavation flow field that develops behind the rapidly expanding shock front, which has moved beyond the boundaries of this illustration. The lines with arrows indicate streamtubes along which material flows downward and outward from the crater. The streamtubes cut across the contours of maximum shock pressure, showing that material ejected at any given range from the impact is shocked to a variety of different maximum pressures. When material flowing through a streamtube crosses the initial surface it forms part of the ejecta curtain. Ejecta emerging near the impact site travel at high speed, whereas ejecta emerging at larger distances travel at slower velocities.

Inside the growing crater, vaporized projectile and target may expand rapidly out of the crater, forming a vapor plume that, if massive enough, may blow aside any surrounding atmosphere and accelerate to high speed. In the impacts of sufficiently large and fast projectiles some of this vapor plume material may even reach escape velocity and leave the planet, incidentally also removing some of the planet's atmosphere. Such "impact erosion" may have played a role in the early history of the Martian atmosphere. Even in smaller impacts the vapor plume may temporarily blow aside the atmosphere, opening the way for widespread ballistic dispersal of melt droplets (tektites) above the atmosphere and perhaps permitting the formation of lunar-like ejection blankets even on planets with dense atmospheres, as has been observed on the Soviet Venera 15/16 images of Venus.

Crater growth rate. The growing crater is at first hemispherical in shape. Its depth $H(t)$ and diameter $D(t)$ both grow approximately as $t^{0.4}$, where t is time after the impact. Hemispherical growth ceases after a time of about $(2H_t/g)^{1/2}$, where H_t is the final depth of the transient crater. At this time the crater depth stops increasing (it may even begin to decrease as collapse begins), but its diameter continues to increase. The crater shape, thus, becomes a shallow bowl, finally attaining a diameter roughly three to four times its depth. At this stage, before collapse modifies it, the crater is known as a "transient" crater. Even simple craters experience some collapse (which produces the breccia lens), so

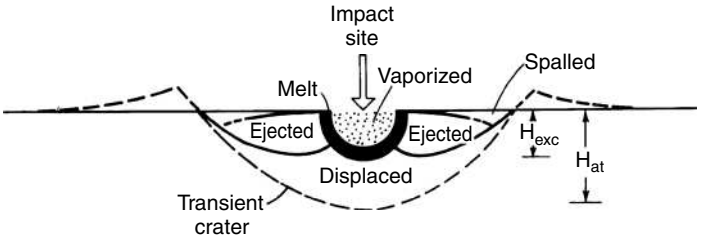

Figure 6.8 Provenance of material expelled from an impact crater. Vaporized material expands outward in a vapor plume. Of the remaining material, some is ejected and some is displaced out of the crater and deforms the adjacent rocks, uplifting the surface near the rim and downwarping rocks beneath the crater floor. The ejected material is excavated from a maximum depth H_{exc} that is only about one-third of the transient crater depth or one-tenth of the transient crater diameter. The dashed lines show the profile of the transient crater.

that the transient crater is always a brief intermediate stage in geological crater formation. However, since most laboratory craters are "frozen" transient craters, much of our knowledge about crater dimensions refers to the transient stage only, and must be modified for application to geological craters.

Maximum depth of excavation. Laboratory, field, and computer studies of impact craters have all confirmed that only material lying above about 1/3 of the transient crater depth (or about 1/10 of the diameter) is thrown out of the crater. Material deeper than this is simply pushed downward into the target, where its volume is accommodated by deformation of the surrounding rocks (Figure 6. 8). Thus, in sharp contrast to ejecta from volcanic craters, material in the ejecta blankets of impact craters does not sample the full depth of rock intersected by the crater, a surprising fact that has led many geologists astray in their estimation of the nature of the ejected debris.

The form of the transient crater produced during the excavation stage may be affected by such factors as obliquity of the impact (although the impact angle must be less than about 6° for a noticeably elliptical crater to form at impact velocities in excess of about 4 km/s), the presence of a water table or layers of different strength, rock structure, joints, or initial topography in the target. Each of these factors produces its own characteristic changes in the simple bowl-shaped transient crater form.

6.3.3 Modification

Shortly after the excavation flow has opened the transient crater and the ejecta has been launched onto ballistic trajectories, a major change takes place in the motion of debris within and beneath the crater. Instead of flowing upward and away from the crater center, the debris comes to a momentary halt, then begins to move downward and back toward the center whence it came. This collapse is generally attributed to gravity, although elastic rebound of the underlying, compressed rock layers may also play a role. The effects of

collapse range from mere debris sliding and drainback in small craters to wholesale alteration of the form of larger craters in which the floors rise, central peaks appear, and the rims sink down into wide zones of stepped terraces. Great mountain rings or wide central pits may appear in still larger craters (Figure 6.9).

These different forms of crater collapse begin almost immediately after formation of the transient crater. The timescale of collapse is similar to that of excavation, occupying an interval of a few times $(D/g)^{1/2}$. Crater collapse and modification, thus, take place on timescales very much shorter than most geologic processes. The crater resulting from this collapse is then subject to the normal geologic processes of gradation, isostatic adjustment, infilling by lavas, etc. on geologic timescales. Such processes may eventually result in the obscuration or even total obliteration of the crater.

The effects of collapse depend on the size of the crater. For transient craters smaller than about 15 km diameter on the Moon, or about 3 km on the Earth, modification entails only collapse of the relatively steep rim of the crater onto its floor. The resulting "simple crater" (see Figure 6.1a) is a shallow bowl-shaped depression with a rim-to-rim diameter D about five times its depth below the rim H. In fresh craters the inner rim stands near the angle of repose, about 30°. Drilling in terrestrial craters (Figure 6.10) shows that the crater floor is underlain by a lens of broken rock (mixed breccia) derived from all of the rock units intersected by the crater. The thickness of this breccia lens is typically 1/2 the depth of the crater H. Volume conservation suggests that this collapse increases the original diameter of the crater by about 15%. The breccia lens often includes layers and lenses of highly shocked material mixed with much less-shocked country rock. A small volume of shocked or melted rock is often found at the bottom of the breccia lens.

Complex craters (Figure 6.1b,c) collapse more spectacularly. Walls slump, the floor is stratigraphically uplifted, central peaks or peak rings rise in the center, and the floor is overlain by a thick layer of highly shocked impact melt (Figure 6.11). The detailed mechanism of collapse is still not fully understood because straightforward use of standard rock mechanics models does not predict the type of collapse observed (see Box 8.1). The current best description of complex crater collapse utilizes a phenomenological strength model in which the material around the crater is approximated as a Bingham fluid, a material that responds elastically up to differential stresses of about 3 MPa, independent of overburden pressure, and then flows as a viscous fluid with viscosity of the order of 1 GPa-s at larger stresses. In a large collapsing crater the walls slump along discrete faults, forming terraces whose widths are controlled by the Bingham strength, and the floor rises, controlled by the viscosity, until the differential stresses fall below the 3 MPa strength limit. A central peak may rise, and then collapse again in large craters, forming the observed internal ring (or rings). Figure 6.9 illustrates this process schematically. The rock in the vicinity of a large impact may display such an unusual flow law because of the locally strong shaking driven by the large amount of seismic energy deposited by the impact.

The mechanics of the collapse that produces multiring basins (Figure 6.1d) is even less well understood. Figure 6.12 illustrates the structure of the Orientale basin on the Moon with a highly vertically exaggerated cross section derived from both geological and

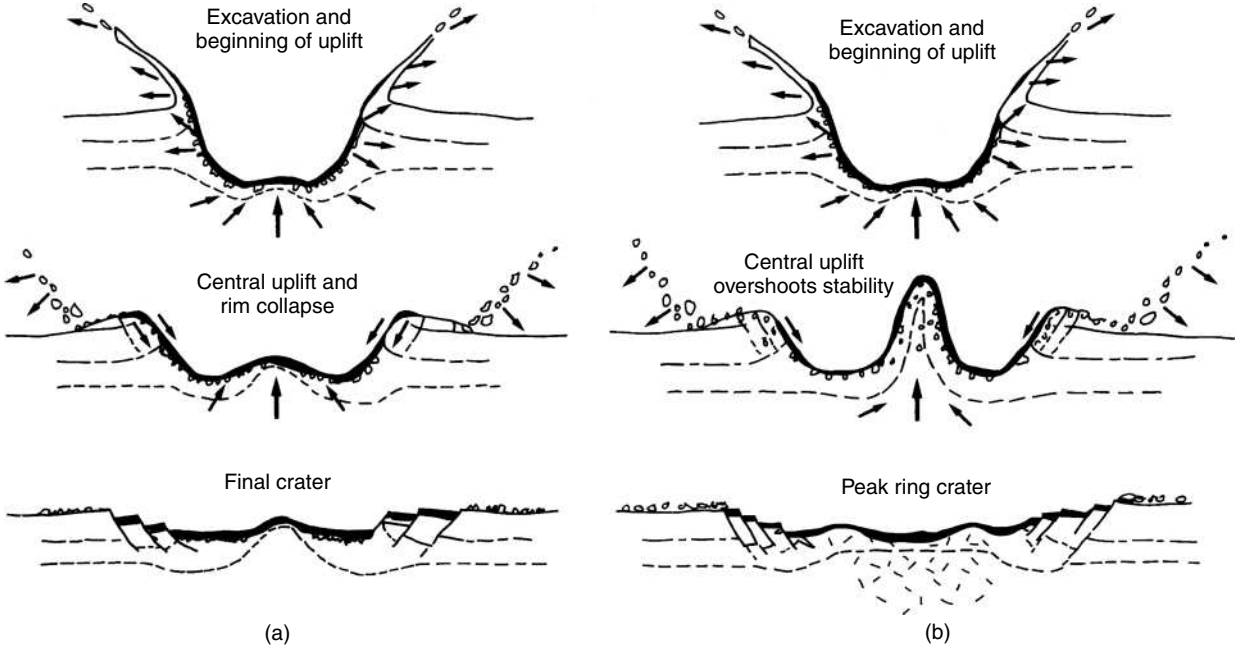

Figure 6.9 Illustration of the formation of complex craters of either (a) central peak morphology or (b) peak ring morphology. Uplift of the crater floor begins even before the rim is fully formed. As the floor rises further, rim collapse creates a wreath of terraces surrounding the crater. In smaller craters the central uplift "freezes" to form a central peak. In larger craters the central peak collapses and creates a peak ring before motion ceases.

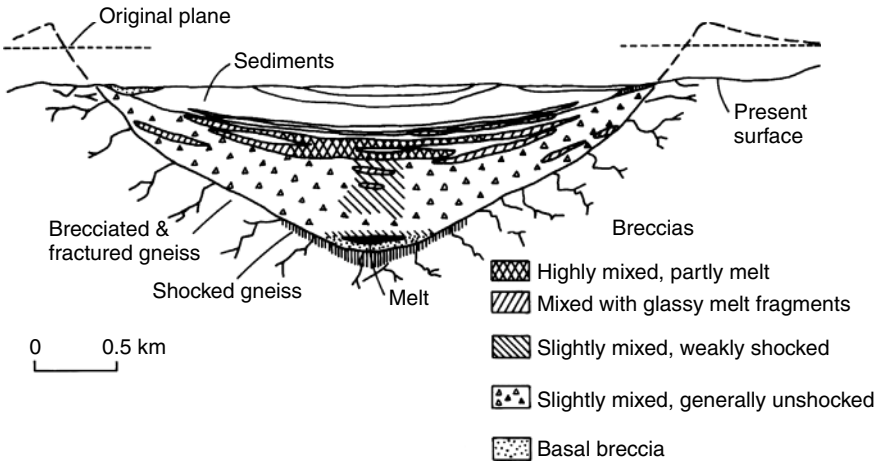

Figure 6.10 Geologic cross section of the 3.4 km diameter Brent Crater in Ontario, Canada. Although the rim has been eroded away, Brent is a typical, simple crater that forms in crystalline rocks. A small melt pool occurs at the bottom of the breccia lens and more highly shocked rocks occur near its top.

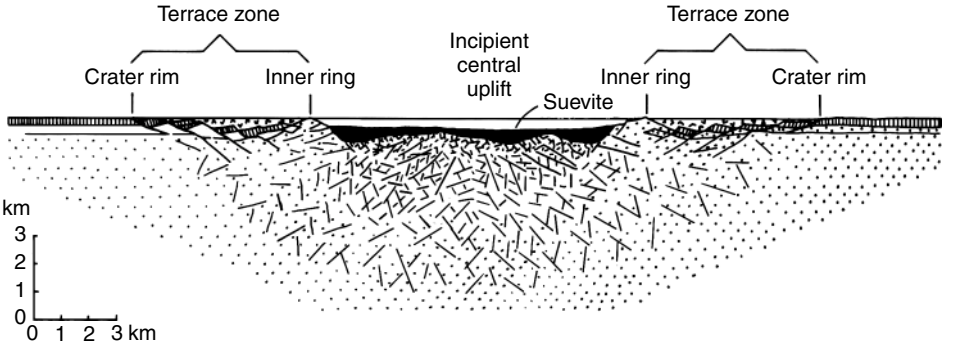

Figure 6.11 Geologic cross section of the 22 km diameter Ries Crater in Germany. Drilling and geophysical data suggest that this is a peak ring crater. Its central basin is filled with suevite, a mixture of highly shocked and melted rocks and cold clasts.

geophysical data. Note the ring scarps are interpreted as inward-dipping faults above a pronounced mantle uplift beneath the basin's center. One idea that is currently gaining ground is that the ring scarps are normal faults that develop as the crust surrounding a large crater is pulled inward by the flow of underlying viscous mantle material toward the crater cavity (Figure 6.13). An important aspect of this flow is that it must be confined in a low-viscosity channel by more viscous material below, otherwise the flow simply uplifts the crater floor, and radial faults instead of ring scarps are the result. Special structural conditions are, thus, needed in the planet for multiring basins to form on its surface, so that a g^{-1} dependence for the transition from complex craters to multiring basins is not expected (or observed). This

Box 6.1 **Maxwell's Z model of crater excavation**

The details of the excavation flow can be determined only by experiment or by elaborate numerical computations. Even such numerical work may have difficulty in correctly computing the final dimensions of the transient crater. However, in 1973 D. Maxwell and K. Seifert proposed a simple analytical model of excavation flow (Maxwell, 1977). This model gives a useful kinematic, although not dynamic, description of the cratering flow field. Like all approximate models, it should not be used to determine fine details.

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

$$u_r = \frac{\alpha(t)}{r^Z} \tag{B6.1.1}$$

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

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

$$u_\theta = (Z-2)\, u_r \frac{\sin\theta}{1+\cos\theta}. \tag{B6.1.2}$$

The geometry of the velocity field defined by this model, seen in Figure 6B.1, is remarkably similar to that computed in both explosion and impact cratering events $Z \simeq 3$. The equation of streamlines in polar coordinates is:

$$r = r_0 (1-\cos\theta)^{\frac{1}{Z-2}} \tag{B6.1.3}$$

where r_0 is a constant that is different for each streamline. It is equal to the radius at which the streamline emerges from the surface ($\theta = 90°$). Taking $r_0 = D_{at}/2$, the radius of the transient crater, the maximum depth of excavation H_{exc} is:

$$H_{\text{exc}} = \frac{D_{at}}{2}(Z-2)(Z-1)^{\frac{1-Z}{Z-2}}. \tag{B6.1.4}$$

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

The total mass ejected from a crater described by the Z model, M_{ej}, is a fraction of the total mass displaced from the transient crater M_{e}:

$$M_{\text{ej}} = \frac{Z-2}{Z-1} M_{\text{e}}. \tag{B6.1.5}$$

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

$$\begin{aligned} u_V &= \alpha / s^Z \\ u_H &= (Z-2)\, u_V \end{aligned} \tag{B6.1.6}$$

Box 6.1 **(cont.)**

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

The *Z*-model presented thus far is a kinematical model useful for describing the form of the excavation flow. Maxwell and Seifert attempted to give it more dynamical content by computing the function $\alpha(t)$ in Equation (B6.1.1). This function gives the strength of the flow at any particular time. Its value is different for each streamtube in the flow. It is estimated by using energy conservation in each of these streamtubes, neglecting interactions with adjacent tubes. Thus, the sum of the kinetic, gravitational, and distortional energies is found in each streamtube at some initial time. The total energy in each streamtube is conserved as the flow progresses. However, the kinetic energy declines at the expense of the gravitational and distortional energies, so that the net flow velocity declines. Unfortunately, this aspect of the *Z* model has not worked out well in practice: The actual course of the excavation flow is best determined through detailed dynamical models.

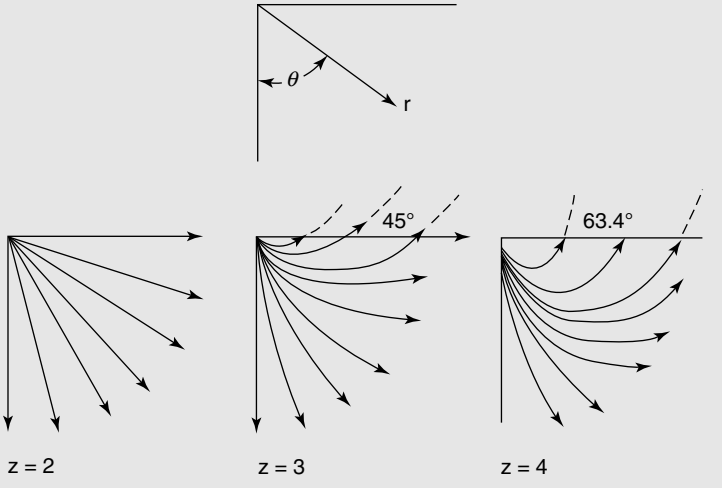

Figure B6.1

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

In spite of all its faults, the *Z*-model gives a reasonably accurate representation of the gross geometric features of the cratering flow and can even be used to predict some first-order dynamical properties. It has the unfortunate feature of not being a truly dynamical model, so that further refinements are not necessarily closer approximations to the full dynamical equations of motion. Nevertheless, the excellent properties of this model are probably still far from being fully exploited.

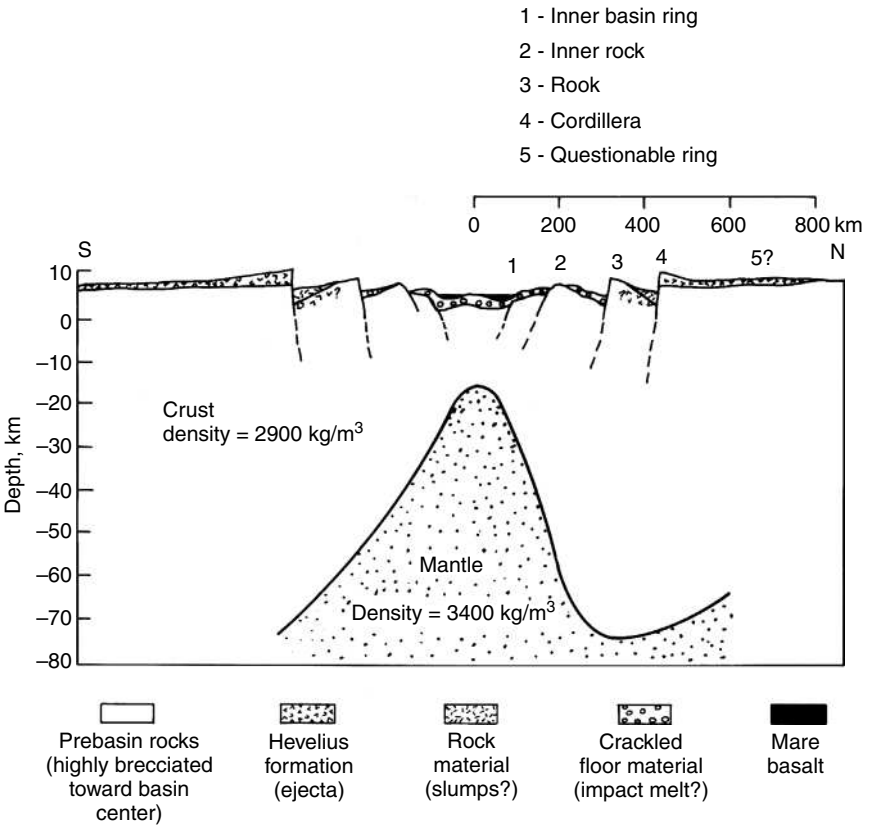

Figure 6.12 Geologic and geophysical structure of the Orientale basin on the Moon, one of the freshest and best-studied multiring basins. A dense mantle uplift underlies the center of the basin. The crustal thinning above the uplift is due to the ejection of about 40 km of crustal material from the crater that formed the basin. The great ring scarps shown in cross-section formed during collapse of the crater. Note the 10X vertical exaggeration necessary to show the ring scarps.

theory is capable of explaining both the lunar-type and Valhalla-type multiring basins as expressions of different lithosphere thicknesses.

6.4 Ejecta deposits

A deposit of debris ejected from the crater interior surrounds essentially all impact craters. The only exceptions are craters on steeply sloping surfaces or on satellites with too little gravity to retain the ejecta or too much porosity to produce it. This ejecta deposit is thickest at the crater rim and thins with increasing distance outwards. Where this deposit is recognizably continuous near the crater it is called an "ejecta blanket." Ejecta beyond the edge of the continuous deposit are thin and patchy. Secondary craters occur in this zone and beyond it. Figure 6.14 shows the ejecta blanket of the 30-km diameter lunar crater Timocharis. The pre-existing

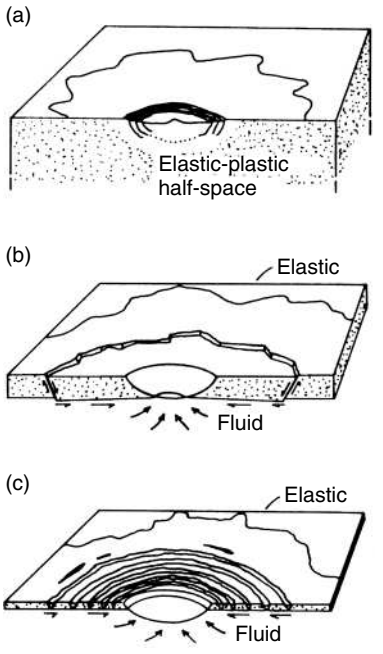

Figure 6.13 The ring tectonic theory of multiring basin formation: (a) shows the formation of a normal complex crater on a planet with uniform rheology; (b) shows the inward-directed flow in a more fluid asthenosphere underlying a lithosphere of thickness comparable to the crater depth and the resulting scarps; (c) shows a Valhalla-type basin developing around a crater formed in a very thin lithosphere.

terrain within about one crater radius of the rim is buried and mostly obliterated by the continuous ejecta blanket. Light patches show where thinner deposits overlie the mare surface.

The thickness of the ejecta deposit varies greatly with direction away from the rim. Azimuthal thickness variations, at the same radius, can be as large as ten to one. The ejecta is concentrated into rays that often are observed to form at angles of about 30° to one another. However, it is also true that, at a given azimuth, the thickness falls rapidly and systematically with distance away from the crater center. Compilation of many data sets from both impact and explosion craters shows that the thickness $\delta(r)$ as a function of radius r away from the crater center is given by an approximate inverse cube relation:

$$\delta(r) = f(R)\left(\frac{r}{R}\right)^{-3\pm0.5} \tag{6.9}$$

where R is the radius of the crater rim. Integration of this relation indicates that most of the mass of the ejecta is located near the crater rim. According to the approximate "Schröter's rule," the volume of the ejecta is approximately equal to the volume of the crater bowl. Although this rule seems to make a great deal of sense (however, it does ignore the increase in volume of the ejected material due to fragmentation), it is unverifiable in practice because the original ground surface can seldom be located with adequate precision.

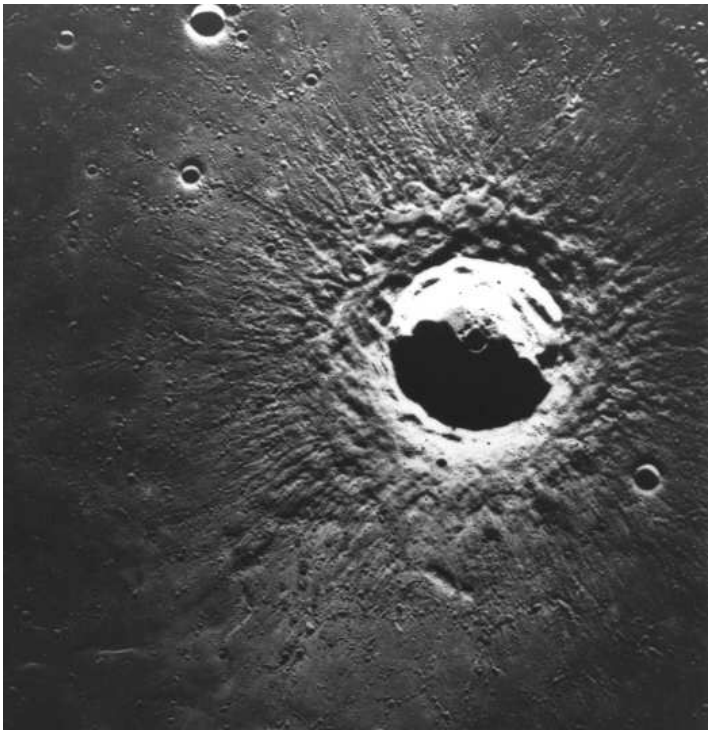

Figure 6.14 Near-vertical view of the 30 km diameter lunar crater Timocharis. The ejecta deposits are dunelike near the crater rim but grade into a subradial facies beyond about 2 *R* from the crater's center. Secondary craters occur at greater distances. A pattern of bright material surrounding the crater indicates the presence of ejecta too thin to greatly modify the pre-existing terrain at this scale. Apollo 15 photo AS15–1005.

The rims of many fresh craters are littered with blocks of rock ejected from beneath the crater. Compilations of data show that block size generally decreases as a function of distance from the rim. The maximum size of block observed on the rim of a crater is related to the size of the crater itself and an empirical relation that holds over a wide range of crater sizes relates the mass of the largest ejected fragment, m_f, to the total mass ejected from the crater, M_e. If both masses are expressed in kilograms, this relation is:

$$m_f = 0.8 M_e^{0.8}. \tag{6.10}$$

6.4.1 Ballistic sedimentation

The ejecta deposited around an impact crater on an airless planet are emplaced ballistically, that is, ejecta are thrown from the crater with some initial velocity, follow a nearly parabolic trajectory above the planet's surface, then fall back to the surface, striking with the

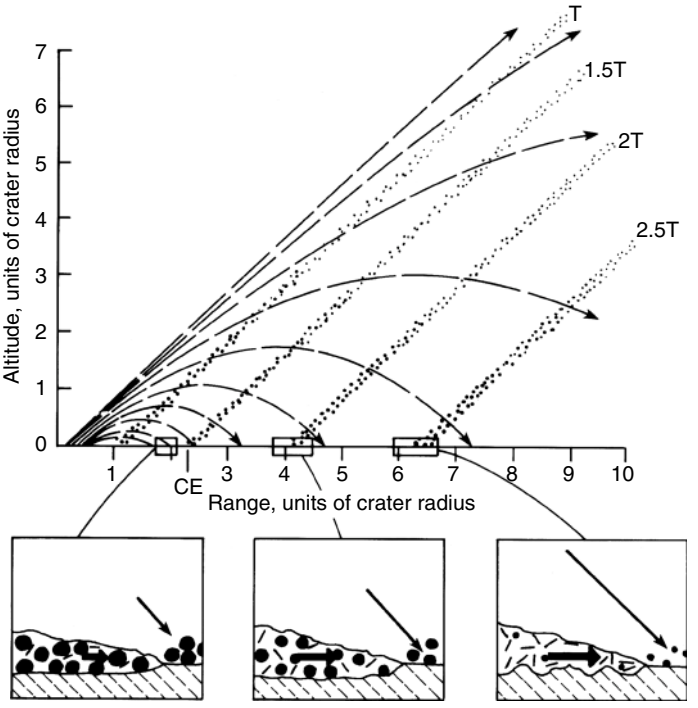

Figure 6.15 The debris ejected from an impact crater follows ballistic trajectories from its launch position within the final crater (the rim of the final crater is located at a range equal to one crater radius). The innermost ejecta are launched first and travel fastest, following the steepest trajectories shown in the figure. Ejecta originating farther from the center are launched later and move more slowly, falling nearer the crater rim. Because of the relationship between position, time, and velocity of ejection, the debris forms an inverted cone that sweeps outward across the target. This debris curtain is shown at four separate times during its flight, at 1, 1.5, 2, and 2.5 T_f, where T_f is the crater formation time, $(D/g)^{1/2}$. Coarser, less-shocked debris travels near the base of the curtain, whereas the fast, highly shocked ejecta fraction tends to travel near the top. The three lower figures show details of the pre-existing ground surface when the ejecta curtain arrives. As the range increases the ejecta strike with progressively larger velocities, incorporating larger amounts of surface material and imparting a larger net horizontal velocity.

same velocity as on ejection. Some interaction may occur between ejecta fragments in the denser parts of the ejecta curtain, but the general motion is dominated by ballistics alone. The ground surface around the crater is profoundly affected by the ejecta as it lands, and its interaction with this falling debris determines the character of the ejecta deposits of large craters.

The debris ejected from an impact crater travels together in the form of an "ejecta curtain." Although each fragment follows a parabolic trajectory, the times and velocities of ejection from the crater are organized so that most of the debris lies on the surface of an expanding inverted cone. Figure 6.15 illustrates the parabolic trajectories of a number

of fragments from an impact crater. The positions of these fragments at several times are indicated. At any one time the debris lies on the surface of a cone that makes an angle of about 45° with the ground surface. This cone sweeps rapidly outward from the crater rim. Debris from the curtain strikes the ground at its base, impacting first near the crater rim, then at greater distances as time progresses. The size of the ejecta fragments near the base of the ejecta curtain is expected to be larger than the fragments higher in the curtain, and the proportion of highly shocked fragments and glass increases with height in the curtain.

Ballistically emplaced debris falling near the crater rim strikes with a low velocity because it travels only a short distance. At the rim itself this velocity is so low that rock units may retain some coherence and produce an overturned flap with inverted stratigraphy. Eugene Shoemaker first recognized such overturned beds at Meteor Crater, Arizona. At greater distances from the crater rim, the debris strikes with a higher velocity. When this velocity is large enough, surface material is eroded and mixes with the debris. The falling ejecta also possess a radially outward velocity component. Although the vertical velocity is cancelled when the debris strikes the surface, the mixture of debris and surface material retains its outward momentum (see the lower inserts in Figure 6.15). This mixture moves rapidly outward as a ground-hugging flow of rock debris, similar in many ways to the flow of a large dry-rock avalanche. Depositional features such as dunes, ridges, and radial troughs indicative of high-speed flow may result from this motion. The deposit itself consists of an intimate mixture of primary crater ejecta and of secondary material scoured from the pre-existing ground surface.

6.4.2 Fluidized ejecta blankets

The ejecta blankets of impact craters on Mars are dramatically different from those on the Moon or Mercury. Martian craters smaller than about 5 km in diameter closely resemble their counterparts on the Moon. However, craters between 5 and 15 km in diameter have a single ejecta sheet that extends about one crater radius from the rim and ends in a low concentric ridge or outward-facing escarpment. These are called “rampart craters,” from the continuous ridge surrounding the ejecta deposit. The ejecta of most craters larger than 15 km in diameter are divided into petal-like lobes that extend two or more radii from the rim (Figure 6.16), approximately twice as far as the continuous ejecta deposits of lunar or Mercurian craters. A few large craters with lunar-type ejecta blankets are known, but they are rare. Many Martian craters have abnormally large central peaks and other internal collapse structures compared with lunar or Mercurian craters, also suggesting the presence of some fluidizing agent peculiar to Mars.

The fluidized ejecta blankets of Martian craters appear to have been emplaced as thin, ground-hugging flows. When impeded by topographic obstacles that could not be overridden, the flows are deflected and either spread out elsewhere or pond against the obstacle. Ejecta lobes fail to overtop low hills and mesas that are only a few times

Figure 6.16 The 19 km diameter crater Yuty on Mars is surrounded by thin, petaloid flow lobes that extend approximately twice as far from the crater as the continuous ejecta deposits of lunar or Mercurian craters. Viking Orbiter frame 3A07.

higher than the flow thickness itself, suggesting that the lobe could not have traveled as dispersed clouds of the base-surge type, nor were they emplaced by ballistic sedimentation, because the ejecta curtain should have fallen on the topographic obstacle from above.

The peculiar form of Martian ejecta blankets is generally attributed to the presence of liquid water in the substrate. Ejected along with subsurface material, liquid water mixed into the ejecta would greatly enhance the mobility of the debris, converting the dry, fragmental ejecta flows characteristic of lunar craters to fluid debris flows similar to terrestrial mudflows. Nevertheless, not all Mars researchers agree with this interpretation and an alternative viewpoint attributes at least part of the mobility of these flows to interaction with the thin atmosphere of Mars.

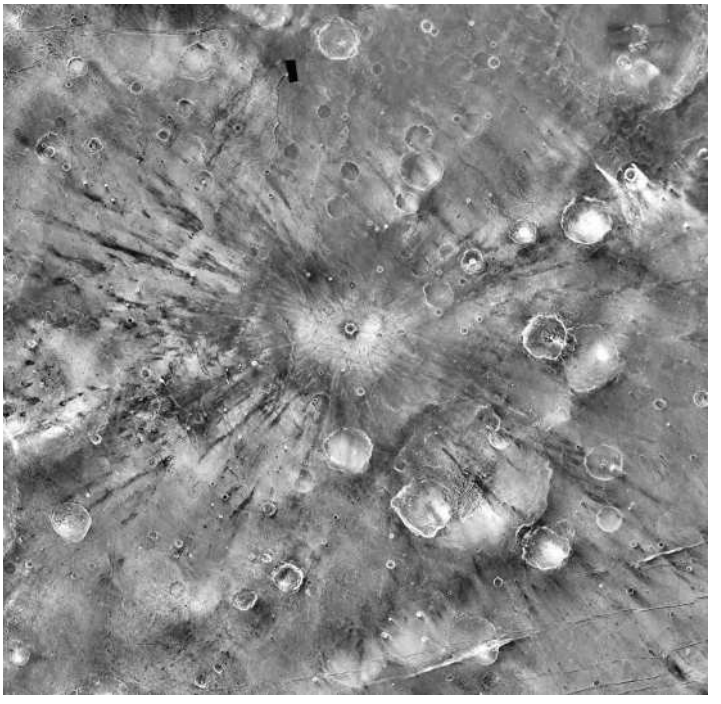

Figure 6.17 Night-time IR thermal image of the 6.9 km diameter Martian crater Gratteri, located at 17.8° S, 202° E. The dark streaks are created by secondary impact craters that extend up to 500 km from the crater center. In images of this type, dark regions are cold and emit little IR radiation because they have low thermal inertia, indicating that the streaks are composed of fine-grained material compared to their surroundings. The overall image measures 545 x 533 km across. THEMIS image courtesy of Phil Christensen. NASA/JPL/ASU. See also color plate section.

6.4.3 Secondary craters

Numerous secondary impact craters, variously occurring either singly or in loops, clusters and lines, surround large impact craters. Figure 6.17 shows the secondary crater field around the crater Gratteri on Mars, as revealed by the THEMIS thermal mapper. Recognizable secondary craters extend from just beyond the continuous ejecta blanket out to distances of up to thousands of kilometers from their source crater. Close to the primary crater, secondary craters are produced by relatively low-velocity impacts and are, thus, irregular in shape, shallow, and obviously clustered, and are often separated by V-shaped dunes known as the "herringbone pattern." Farther from the primary impact, velocities are larger and secondary craters are more dispersed, which makes them difficult to discriminate from small primary craters.

An important controversy is presently raging about the importance of secondary craters in masking the primary flux. If a majority of the small craters (less than a few 100 m diameter) on a planet's surface are secondary, then ages assigned to cratered surfaces based on the

assumption that the craters are primary will be too great. On the other hand, many experienced crater counters claim that they can exclude secondary craters because they are clustered, a claim that is disputed by other experts. At the moment there is no consensus on this problem.

The maximum size of secondary craters is approximately 4% of the primary diameter on the Moon, Mars, and elsewhere. However, on Mercury obvious secondary craters are apparently several times larger. This observation is at odds with the larger impact velocity on Mercury, which is expected to result in smaller ejected fragments, not larger ones. Is the crust of Mercury somehow stronger than that of the Moon or Mars? At the moment, the solution to this conundrum is unknown.

6.4.4 Oblique impact

Although high-velocity impact craters are circular down to very low angles of approach, the pattern of the ejecta may betray impact obliquity at angles as large as 45°. The first sign of an oblique impact is an asymmetric, but still bilaterally symmetrical, ejecta blanket. The ejecta in the uprange direction are thinner and less extensive than those in the downrange direction at impact angles near 45°. At impact angles near 30° an uprange wedge free of ejecta develops, an example of which is shown in Figure 6.3. As the angle decreases still farther, to 10°, ejecta-free regions appear in both downrange and uprange directions, although bright streaks may extend downrange in very fresh craters. At such low angles the crater itself becomes elliptical, with its long axis parallel to the flight direction of the projectile. At these highly oblique angles the projectile essentially plows a furrow into the target surface, throwing ejecta out to both sides to form a "butterfly-wing" pattern.

6.5 Scaling of crater dimensions

One of the most frequently asked questions about an impact crater is, "How big was the meteorite that made the crater?" Like many simple questions this has no simple answer. It should be obvious that the crater size depends upon the meteorite's speed, size, and angle of entry. It also depends on such factors as the meteorite's composition, the material and composition of the target, surface gravity, presence or absence of an atmosphere, etc. The question of the original size of the meteorite is usually unanswerable, because the speed and angle of impact are seldom known. The inverse question, of how large a crater will be produced by a given-sized meteorite with known speed and incidence angle is in principle much simpler to answer. However, even this prediction is uncertain because there is no observational or experimental data on the formative conditions of impact craters larger than a few tens of meters in diameter, while the impact structures of geologic interest range up to 1000 km in diameter. The traditional escape from this difficulty is to extrapolate beyond experimental knowledge by means of scaling laws.

C. W. Lampson, who studied the craters produced by TNT explosions of different sizes, introduced the first scaling law in 1950. Lampson found that the craters were similar to

one another if all dimensions (depth, diameter, depth of charge placement) were divided by the cube root of the explosive energy W. Thus, if the diameter D of a crater produced by an explosive energy W is wanted, it can be computed from the diameter D_0 of a crater produced by a smaller explosive energy W_0 using the proportion:

$$\frac{D}{D_0} = \left(\frac{W}{W_0}\right)^{1/3}. \tag{6.11}$$

An exactly similar proportion may be written for the crater depth, H. This means that the ratio of depth to diameter, H/D, is independent of yield, a prediction that agrees reasonably well with observation. In more recent work on large explosions the exponent 1/3 in this equation has been modified to 1/3.4 to account for the effects of gravity on crater formation.

Although impacts and explosions have many similarities, a number of factors make them difficult to compare in detail. Thus, explosion craters are very sensitive to the charge's depth of burial. Although this quantity is well defined for explosions, there is no simple analog for impact craters. Similarly, the angle of impact has no analog for explosions. Nevertheless, energy-based scaling laws were very popular in the older impact literature, perhaps partly because nothing better existed, and many empirical schemes were devised to adapt the well-established explosion scaling laws to impacts.

6.5.1 Crater diameter scaling

This situation has changed rapidly in the last few decades, however, thanks to more impact cratering experiments specifically designed to test scaling laws. It has been shown that the great expansion of the crater during excavation tends to decouple the parameters describing the final crater from the parameters describing the projectile. If these sets of parameters are related by a single, dimensional "coupling parameter" (as seems to be the case), then it can be shown that crater parameters and projectile parameters are related by power-law scaling expressions with constant coefficients and exponents. Although this is a somewhat complex and rapidly changing subject, the best current scaling relation for impact craters forming in competent rock (low-porosity) targets whose growth is limited by gravity rather than target strength (i.e. all craters larger than a few kilometers in diameter) is given by:

$$D_{tc} = 1.161\left(\frac{\rho_p}{\rho_t}\right)^{1/3} L^{0.78}\, v_i^{0.44}\, g^{-0.22} \sin^{1/3}\theta \tag{6.12}$$

where D_{tc} is the diameter of the transient crater at the level of the original ground surface, ρ_p and ρ_t are densities of the projectile and target, respectively, g is surface gravity, L is projectile diameter, v_i is impact velocity and θ is the angle of impact from the horizontal. All quantities are in SI units.

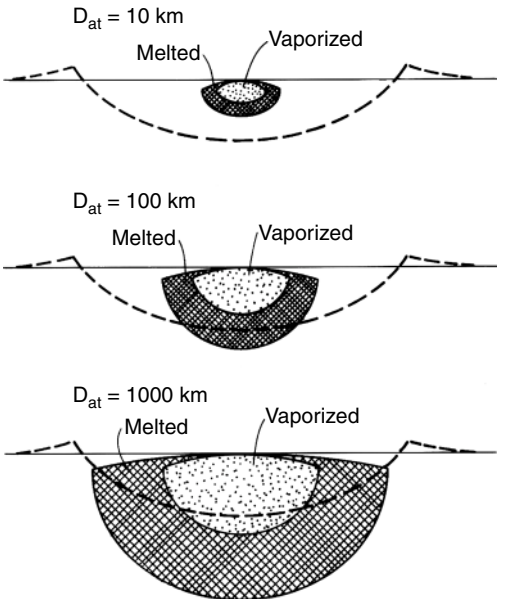

Figure 6.18 The different scaling laws for crater diameter and melt or vapor volume imply that as the crater diameter increases, the volume of melted or vaporized material may approach the volume of the crater itself. This figure is constructed for impacts at 35 km/s on the Earth.

The transient crater depth, H_{tc}, appears to be a constant times the diameter D_{tc}. Although a few investigations have reported a weak velocity dependence for this ratio, the experimental situation is not yet clear.

6.5.2 Impact melt mass

The amount of material melted or vaporized by an impact is a strong function of the impact velocity itself. The melt mass depends principally upon the velocity and mass of the projectile, but does not depend upon the gravitational acceleration of the target body. Very little melt is produced until a threshold velocity of about 12 km/s is reached. Once this threshold is exceeded, the mass of melt M_m is given in terms of the mass of the projectile M_p as:

$$\frac{M_m}{M_p} = 0.25\frac{v_i^2}{\varepsilon_m}\sin\theta, \qquad v_i \geq 12 \text{ km/s} \tag{6.13}$$

where ε_m is the specific internal energy of the Rankine–Hugoniot state from which isentropic decompression ends at the 1 bar point on the liquidus. It is equal to 5.2 MJ/kg for granite, which can be taken to be representative of crustal rocks.

Although the mass of melt does not depend on the gravitational acceleration of the planet, the crater size does, through Equation (6.12). The volume of the melt relative to the volume

of the crater is relatively small for small craters, but as the crater size increases it becomes a progressively larger fraction of the total crater volume, as shown in Figure 6.18. At some sufficiently large diameter (about 1000 km on the Earth), the volume of the melt equals the volume of the crater itself and substantial changes in the morphology of the crater can be expected, although these changes are not well understood at the present time.

6.6 Atmospheric interactions

As fast-moving meteoroids enter the atmosphere of a planet, they are slowed by friction with the atmospheric gases and compressed by the deceleration. Small meteoroids are often vaporized by frictional heating and never reach the surface of the planet. Larger meteoroids are decelerated to terminal velocity and fall relatively gently to the surface of the planet. The diameter of a meteorite that loses 90% of its initial velocity in the atmosphere is typically about 1 m for Earth and 60 m for Venus. However, this assumes that the projectile reaches the surface intact, whereas in fact aerodynamic stresses may crush all but the strongest meteorites. Once fractured, the fragments of an incoming meteorite travel slightly separate paths and strike the surface some distance apart from one another. This phenomenon gives rise to the widely observed strewn fields of meteorites or craters on the Earth, which average roughly a kilometer or two in diameter. On Venus, clusters of small craters attributed to atmospheric breakup are spread over areas roughly 20 km in diameter. Rather surprisingly, clusters of small craters are also observed on Mars, where the spread of small craters in a cluster averages only a few tens of meters to hundreds of meters across.

The aerodynamic crushing stress experienced by an incoming meteoroid is of the order of the stagnation pressure, given by

$$P_{\text{stagnation}} \approx \rho_a v^2 \tag{6.14}$$

where ρ_a is the density of the atmosphere through which the meteoroid is traveling and v is the relative velocity of the meteoroid and atmosphere. Evidently, even the thin Martian atmosphere is enough to fracture and partially disperse weak incoming meteoroids.

It can be shown that the dispersion of a cluster of fragments is a maximum when breakup occurs at twice the atmospheric scale height, H_s. In this case the expected dispersion ΔY is given by

$$\Delta Y \approx \frac{2H_s}{\sin\theta}\sqrt{\frac{\rho_a}{\rho_p}} \tag{6.15}$$

where θ is the angle of entry of the meteoroid with respect to the horizontal and ρ_p is its density. Because the scale heights of the atmospheres of the Earth, Venus, and Mars are all similar, the dispersion of clusters of fragments is expected to be roughly a factor of ten different among these bodies, increasing from Mars to Earth to Venus, as observed.

The atmospheric blast wave and thermal radiation produced by an entering meteorite may also affect the surface: The 1908 explosion at Tunguska River, Siberia, was probably

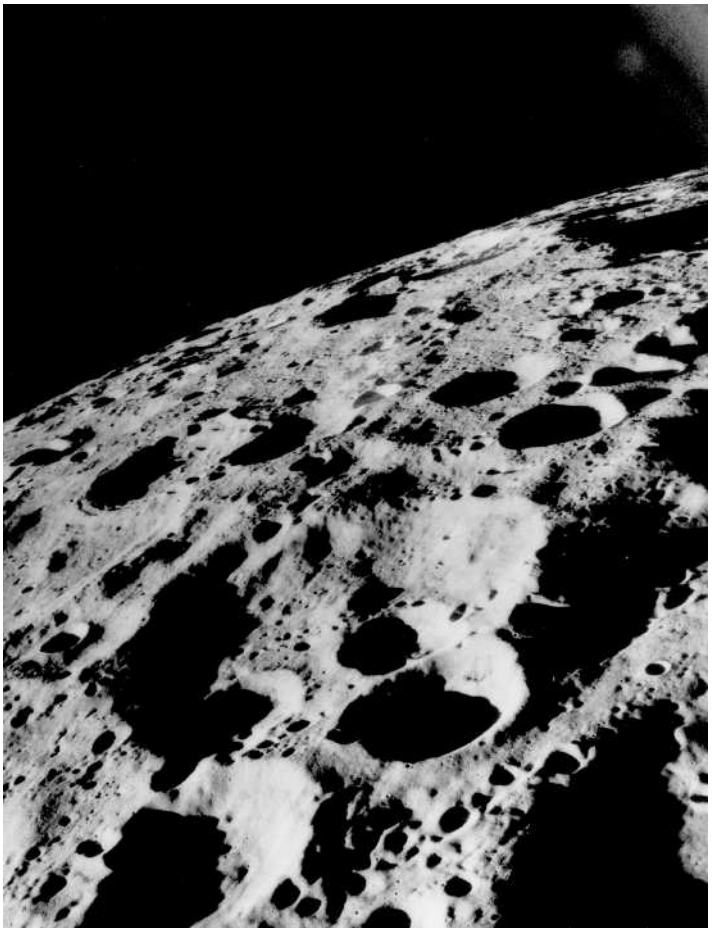

Figure 6.19 Oblique view of a heavily cratered landscape on the Moon. This area is to the northeast of crater Tsiolkovskiy on the Moon's farside. The large crater near the center is about 75 km in diameter, but craters as small as a few tens of meters in diameter can be discerned in the foreground. Apollo 17 photo 155–23702 (H).

produced by the entry and dispersion of a 30 to 50 m diameter stony meteorite that leveled and scorched about 2000 km^2 of meter-diameter trees. Radar-dark "splotches" up to 50 km in diameter on the surface of Venus are attributed to pulverization of surface rocks by strong blast waves from meteorites that were fragmented and dispersed in the dense atmosphere.

6.7 Cratered landscapes

Impact craters have been treated as individual entities in the preceding sections. However, as spacecraft images abundantly illustrate, the surfaces of most planets and satellites are

scarred by vast numbers of impact craters that range in size from the limit of resolution to a substantial fraction of the planet's or satellite's radius. In some places impact craters are the dominant landform: little of the observed topography can be ascribed to any other process (see Figure 6.19). Craters on such a surface exhibit degrees of preservation ranging from fresh craters with crisp rims and bright rays to heavily battered or buried craters that may only betray their presence by a broken rim segment or a ragged ring of peaks.

The present crater population on surfaces such as the lunar highlands or the more lightly cratered lunar mare is the outcome of a long history of impact cratering events. Analysis of the existing crater population in conjunction with some assumptions about the rate of crater formation may reveal a great deal about the geologic history of a surface. A typical population is composed of craters with a wide range of sizes, some of which are relatively fresh, with sharp rims, extensive rays, and crisp fields of small secondary craters, whereas others are progressively more degraded. On parts of such airless bodies as the Moon, the principal agent of degradation is other impacts, producing a surface that appears to be crowded with craters. In other regions volcanism has created plains that are more or less sparsely cratered. Elsewhere in the Solar System the activities of wind, water, or tectonic processes such as subduction erase craters within a short time of their formation, leading to landscapes like the Earth's, where impact craters are among the rarest of landforms, or like that of Venus, where the low abundance of craters may be due to an ancient era of resurfacing that obliterated most pre-existing craters.

Study of crater populations is, thus, a powerful tool for geologic investigation of the surfaces of other planets and satellites. If the flux and size distribution of the impacting bodies were known, studies of crater populations could yield absolute ages of the surface and some of its features. Although the original flux is often unknown, relative ages can usually be obtained. Before geologic inferences can be drawn from crater populations, however, we must have an effective means of describing and comparing them. Unfortunately, a large number of descriptions have evolved over the years as each group of scientists studying a particular problem created their own specialized means of presenting population data, making it difficult to compare the results of different groups. In this section I adopt the major recommendations of a NASA panel convened in 1978 (NASA, 1978) to standardize the presentation of crater population data.

6.7.1 Description of crater populations

The first step in an investigation of the crater population on a given surface is to select an area that is believed to have had a homogeneous geologic history. It would make little sense, for example, to combine the crater population of a sparsely cratered lava plain with that of a densely cratered upland. Once such an area is selected, the craters that lie within it are counted. Most crater population studies include all recognizable craters, regardless of their state of degradation. Where a large enough population exists, more specialized studies may be performed in which the numbers of fresh craters, slightly degraded craters, degraded craters, etc., are counted separately. Although these studies leave some room for

interpretation as to what a "recognizable" or "fresh" crater is, intercomparison of results between different groups of crater counters has generally shown good agreement.

Craters occur in a wide variety of sizes, so that the principal information about a crater population is the number of craters per unit area as a function of crater diameter. It is presumed that impact cratering is a random process and that there is no significance to the particular location of craters within the selected area, so that only data on the number and diameter of craters is kept.

Incremental distribution. Numerous ways of representing the number of craters as a function of diameter have been developed. One very simple method is to list the number N of craters per unit area with diameters between two limits, say between D_a and D_b. The problem with this method is that the resulting number of craters depends upon the interval $\Delta D = D_b - D_a$, and different crater counters may choose different intervals. Furthermore, if the interval ΔD is fixed at, say, 1 km, this might be convenient for craters with diameters between 5 and 20 km, but would be too large for craters with diameters less than 1 km and too small for craters larger than 100 km. A simple way to overcome this problem is to let the interval depend on crater size. Thus, the number of craters may be tabulated between D and $2D$, where the intervals increase in octaves. Actually, this binning has been found to be too coarse in practice, so that most such *incremental* size-frequency distributions use an interval of D to $\sqrt{2}D$. The incremental distribution still suffers, however, from the arbitrary choice of a starting diameter D. It is now recommended that the bins be chosen so that one bin boundary is at $D = 1$ km.

Cumulative distribution. Although the incremental size-frequency distribution could be successful if the same bin sizes and boundaries are universally adopted, it lacks fundamental simplicity. Another distribution has long been used that is independent of bin size: this is the *cumulative* size-frequency distribution. In this distribution the number N_{cum} of craters per unit area with diameters greater than or equal to a given diameter D is tabulated. Not only is the resulting distribution $N_{\text{cum}}(D)$ independent of bin size, but any desired incremental distribution can be easily generated from it, since the number of craters N per unit area in the interval between D_a and D_b is simply:

$$N(D_a, D_b) = N_{\text{cum}}(D_a) - N_{\text{cum}}(D_b) \tag{6.16}$$

where $D_b > D_a$. $N(D_a,D_b)$ is necessarily positive or zero by the definition of the cumulative number distribution. The only disadvantage of the cumulative distribution is that the cumulative number of craters at some given diameter depends upon the number of craters at all larger diameters. Although this is rarely a major problem, cumulative distributions in limited-diameter intervals (often controlled by the size of the region being analyzed) have to be adjusted in overall value to join with the distributions from other diameter ranges. The slope of the cumulative plot is not, of course, affected by such adjustments.

It has been found in practice that the cumulative number distribution closely approximates a power function of diameter:

$$N_{\text{cum}} = c\, D^{-b} \tag{6.17}$$

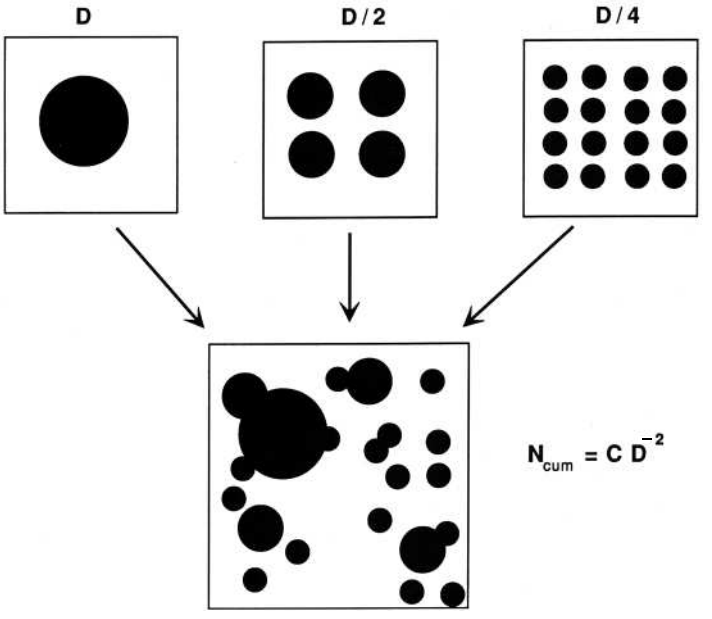

Figure 6.20 A population of craters with a slope of $b = 2$ is structured such that for each crater of diameter D, there are four times as many craters of diameter $D/2$ and 16 times as many as for diameter $D/4$. When mixed together, the population looks very much like that seen on the heavily cratered areas of the Moon and other planets. Note that the total area in each size class of this distribution is the same, a characteristic that may explain why such distributions are so common when formed by either coagulation or impact fragmentation – processes that both depend on surface area.

where $b = 1.8$ for post-mare craters on the Moon between 4 and several hundred kilometers in diameter (and, within the limited data, also seems to hold for impact craters on the Earth).

It is intriguing that the power b is close to 2. A power 2 in Equation (6.17) is a kind of "magic" number because when $b = 2$ the coefficient c is dimensionless (remember N_{cum} is number per unit area). There is no fundamental length or size scale in a crater population with this power law, so that such a population looks the same at all resolutions (see Figure 6.20 for an illustration of such a distribution). It is impossible to tell from a photograph of such a cratered landscape whether the scale of the photograph is 100 km or 1 m (of course, other clues than crater population alone may give a hint about the actual scale). A population of craters described by a power near 2 might arise either from a simple formation process in which there is truly no fundamental length scale or from a series of independent processes that are so complex and chaotic that no one scale dominates.

Because the cumulative number distribution of Equation (6.17) falls as a power of D, it is conventional to graph such distributions on a log-log plot on which a power law is a straight line with slope equal to $-b$ (Figure 6.21). Unfortunately, cumulative number distributions plotted in this way have a tendency to look all the same, apparently differing only in the absolute number density of craters. Although this may be adequate or even desirable for some purposes,

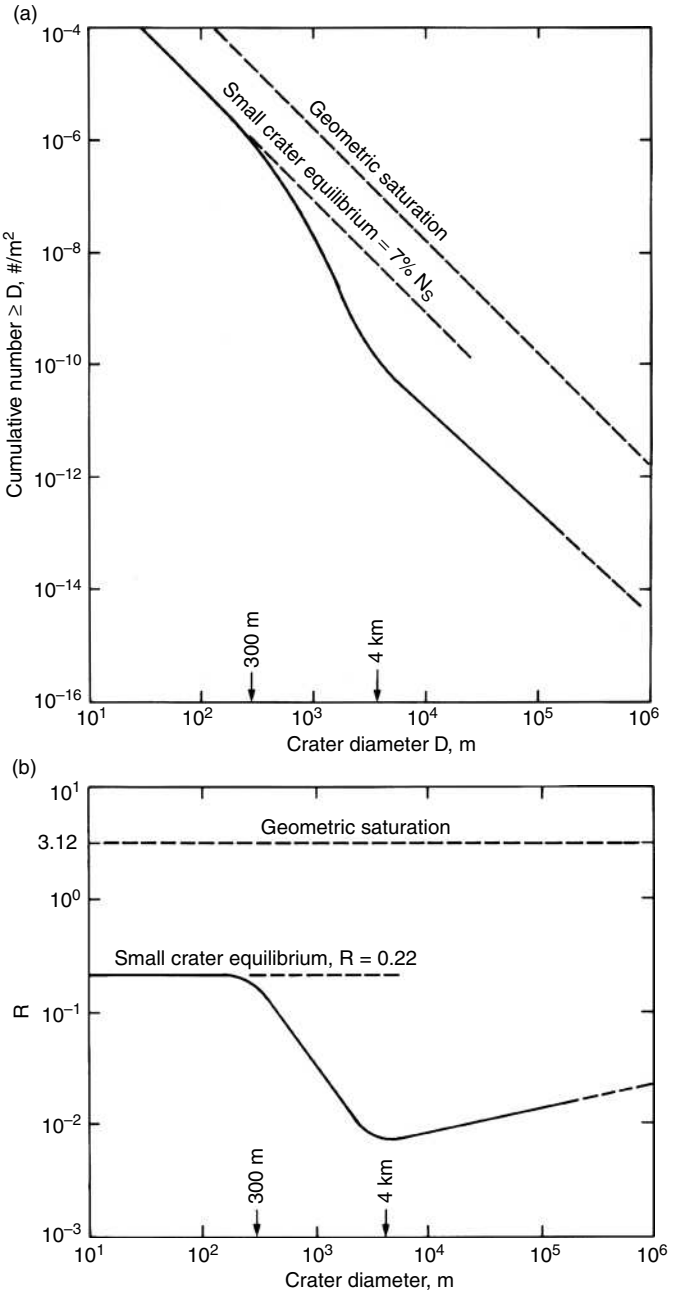

Figure 6.21 The post-mare crater population on the Moon in both a cumulative plot (a) and R-plot (b) format. The population shows three distinct segments. At small diameters ($D < 300$ m) the population is in equilibrium and the cumulative number is proportional to D^{-2}. R is constant for this population. At intermediate diameters (300 m < D < 4 km) the cumulative number is proportional to $D^{-3.4}$, whereas at larger diameters ($D > 4$ km) it falls as $D^{-1.8}$. The curves are dashed at the largest diameters ($D > 200$ km) because no craters of this size have yet formed on the mare. The changes in slope of the crater population are especially evident in the R-plot format. The dashed line labeled "geometric saturation" is an upper limit to the crater density on any surface. After Melosh (1989, Figure 10.2).

such a mode of presentation obscures slight but significant differences in crater populations. Since these differences can be useful in deciphering the source of the crater population or other geologic processes that acted on it, another type of plot is in common use.

The R plot. This type of plot exploits the close approach of b to 2 in Equation (6.17) by graphing essentially the ratio between the actual crater distribution and a distribution with slope –2. A crater population in which the actual slope is −2 would, thus, plot as a horizontal line. The conventional plot of this type is called an R plot (R stands for "Relative"). It is based on an incremental distribution with $\sqrt{2}$ intervals between diameter bins. Note that the slope of this type of incremental plot is the same as the cumulative distribution of Equation (6.17), although the coefficient is different. It is easy to show that if b is constant over the interval D to $\sqrt{2}D$, then the incremental number density is:

$$N(D,\sqrt{2}D) = c\,(1-2^{-b/2})\,D^{-b}. \tag{6.18}$$

The definition of the R plot includes several numerical factors. In terms of the cumulative number distribution it is given by:

$$\mathrm{R}(D) \equiv \frac{2^{3/4}}{\sqrt{2}-1} D^2 \left[N_{\mathrm{cum}}(D) - N_{\mathrm{cum}}(\sqrt{2}D) \right]. \tag{6.19}$$

Figure 6.21 compares the cumulative and R plots for pose-mare craters on the Moon. A useful interpretation of R is to note that, up to a factor of 3.65, R is equal to the fraction $f_c(D)$ of the total area covered by craters in the diameter interval D to $\sqrt{2}D$:

$$\mathrm{R}(D) = 3.65\, f_c(D) \tag{6.20}$$

where both R and f_c are dimensionless numbers. With this interpretation it is easy to see that in a crater population with $b = 2$, for which R and f_c are constant, craters in every size interval occupy the same fraction of the total area. If $b < 2$, as it is for post-mare craters on the moon, R(D) increases as D increases so that large craters occupy a larger fraction of the surface than small craters. If $b > 2$, small craters occupy a larger fraction of the surface than do large ones.

The production population. The impact of the primary meteoroid flux on a planetary surface results in some definite rate of crater production as a function of diameter. As the surface, initially taken to be craterless, ages, more and more craters accumulate on it. The integral of the crater-production rate over the age of the surface is a special, theoretical, crater population called the production population. The production population is the size-frequency distribution of all the craters, excluding secondary craters, that have ever formed since craters began to accumulate on the surface. The population is theoretical in the sense that it neglects all crater-obliteration processes and is, hence, formally unobservable, although the crater population on lightly cratered surfaces may approach the production population closely enough for practical applications.

The production population is a useful concept for the study of the evolution of crater populations. Such studies usually begin with an assumed or inferred production population and

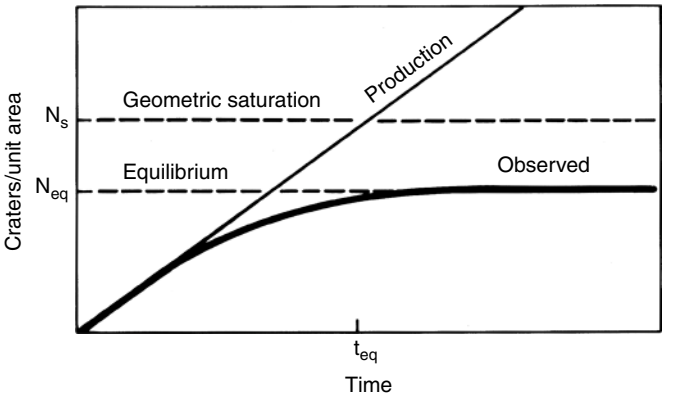

Figure 6.22 Evolution of a crater population in which all craters are the same size. The graph shows how the crater density increases as a function of time. Although the production population rises linearly with time, the number of craters that can be counted on the surface eventually reaches a limit well below the geometric saturation limit. Once the population has reached equilibrium, each additional crater obliterates, on average, one old crater.

then postulate a model of crater obliteration that will, it is hoped, result in a predicted crater population that matches the observed population. The crater-obliteration model is a function of the process being modeled. In the next section we shall consider only the process of crater obliteration by other craters. As we shall see, this can be a surprisingly complex process.

6.7.2 Evolution of crater populations

Given the fact that craters accumulate randomly on the surface of a planet at a rate that is, on average, constant, it should be possible to begin with a knowledge of the rate at which craters form as a function of their size and then predict the crater population at any future time. Furthermore, it should be possible, within certain limits, to invert an evolved crater population and deduce both how long it has been accumulating (that is, the age of the surface) and the rate of crater formation as a function of size.

Simple as these propositions may seem, there has been great difficulty in actually implementing them. The main problem is the interaction between craters of different sizes: the formation of a single large crater on a surface may obliterate many smaller craters, while it takes many small craters to batter a large crater beyond recognition. It has taken many years to fully understand the effects of this interaction on crater populations, and some aspects of it are not completely understood today.

The conceptually simplest population is one in which all craters are the same diameter. Although no natural examples of such a population are known, the study of its evolution introduces several important concepts. Moreover, there are crater populations, such as that on Mimas, in which many of the craters fall within a relatively narrow size range and that may, therefore, be approximated by a population of craters all of the same size.

The equilibrium population. The evolution of this crater population is illustrated in Figure 6.22 as a function of time. The plot shows that the observed crater population and the production population initially increase at the same rate. However, as the density of craters on the surface increases a few older craters are either overlapped by new ones or are buried by their ejecta. As this process continues, some older craters are completely obliterated by younger craters and the observed crater density falls below the production line. Eventually, the crater density becomes so high that each new crater that forms obliterates, on average, one older crater. At this stage the crater population has reached *equilibrium*: no further increase in crater density is possible, although new craters continue to form and the production population increases steadily in number.

The attainment of equilibrium places severe constraints on attempts to date planetary surfaces by crater counting. Up until equilibrium is attained at time t_{eq} in Figure 6.22 the crater density increases with the age of the surface, so that knowledge of the crater production rate permits computation of an absolute age from the crater density. Even if the production rate is not known, the relative ages of two surfaces may be obtained by comparing their crater densities. However, once equilibrium is attained the crater density becomes constant and only a lower limit on the age can be obtained. The relative ages of two surfaces in equilibrium are also completely unconstrained, since their crater densities are identical even though the surfaces may be widely different in age.

Geometric saturation. A useful concept introduced by Don Gault in 1981 (Project, 1981) is that of *geometric saturation*. The idea is to define a crater density that serves as an upper limit to the number of craters that can possibly be recognized on a heavily cratered surface. For a population of craters all of the same size this limit is simply the number of craters per unit area in a hexagonally closest packed configuration, neglecting any possibility of obliteration by overlapping ejecta blankets. In this case, $N_s = 1.15\ D^{-2}$. The definition of a limiting crater density when craters are of different sizes is less objective. Gault proposed the limit:

$$N_{cs} = 1.54\ D^{-2}. \tag{6.21}$$

This crater density corresponds to R = 3.12 or a fractional area coverage $f_c = 0.85$. These limits are shown in Figure 6.21: It is clear that lunar crater populations fall well short of this limiting density, although the mare surface is apparently in equilibrium for crater diameters less than about 300 m. The crater density on even the most heavily cratered surfaces seldom reaches more than about 3–5% of the geometric saturation limit, Equation (6.21).

6.8 Dating planetary surfaces with impact craters

It has long been recognized that the number of impact craters per unit area could date planetary surfaces. Shoemaker *et al.* (1963), who wrote the classic paper on this subject, realized that relative dates on the same planet could always be attained, but absolute dates require knowledge of the flux of impacting bodies.

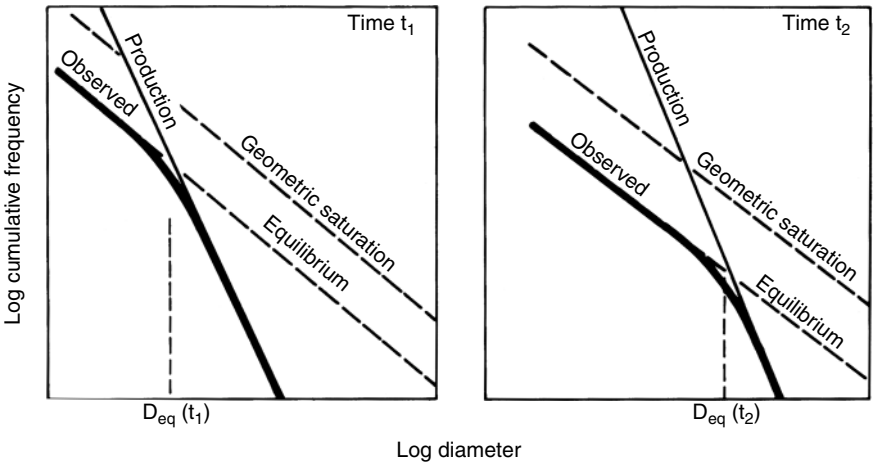

Figure 6.23 Evolution of a crater population with slope $b > 2$. The production population exceeds the equilibrium line at small crater diameters. Small craters are, thus, in equilibrium up to some diameter D_{eq}, above which the observed population follows the production population. The left panel illustrates the population at a relatively early time t_1 and the right panel shows how the population has changed at a later time t_2. The equilibrium diameter D_{eq} increases as a function of time, although this increase is generally not linear.

A full understanding of how to interpret crater populations as a function of size has been long in coming. The attainment of equilibrium by crater populations divides into two distinct cases (and a trivial intermediate). The first, and simplest, case occurs when the production population has a slope b steeper than 2. This case was studied by Gault (1970) and offers the fewest conceptual difficulties. Unfortunately, this distribution is only appropriate for small (≤300 m) craters on the lunar mare (see Figure 6.21), although at the time Gault performed his analysis it was believed to be valid for all crater sizes. The size-frequency distribution of larger lunar craters follows a power law with a slope $b = 1.8$. This distribution fits the second case of an evolving population with slope b smaller than 2, and will be treated shortly.

6.8.1 $b > 2$ population evolution

Figure 6.23 illustrates the evolution of a production population with slope b steeper than 2. The left frame depicts the population at an early time t_1 and the right frame is at a later time t_2. Because the production population is steeper than the geometric saturation line, mutual obliteration must occur for sufficiently small craters (craters smaller than $D_{eq}(t_1)$ in Figure 6.23) no matter how early the time, unless, of course, the population has had so little time to evolve that the statistics of small numbers of craters begins to play a role. Because the actual crater density cannot reach the geometric saturation limit, the observable crater density reaches equilibrium somewhere below this line. The difference in the

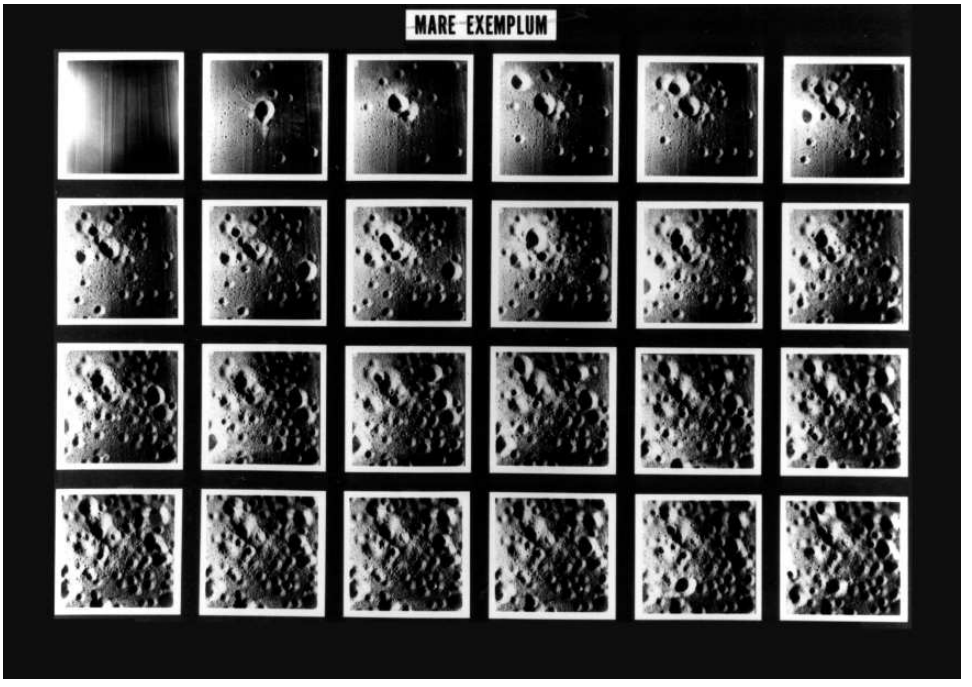

Figure 6.24 A laboratory-scale demonstration of the concept of crater equilibrium. The photographs are of a box 2.5 m square filled 30 cm deep with quartz sand. The sand is topped with 2 cm of carborundum powder to provide a color contrast. Six sizes of projectile were fired into the box at random locations, simulating a production population with a slope index $b = 3.3$, similar to that of small craters on the Moon. Time increases from the upper left horizontally to lower right. Equilibrium is attained about halfway through the simulation: Although individual surface details vary from frame to frame, the crater population in the later frames remains the same. From Gault (1970); photo courtesy of R. Greeley.

number of craters between the projected production population and the observed population is equal to the number of craters obliterated by later impacts. It seems intuitively reasonable that the equilibrium population should follow a line parallel to, but below, the geometric saturation distribution. Gault showed empirically in small-scale impact experiments (Figure 6.24) that this is the case, and subsequent work has confirmed this result both analytically (Soderblom, 1970) and by Monte Carlo computer simulations (Woronow, 1977). In Gault's experiments the crater equilibrium density depends on the slope b of the production population, with steeper slopes giving a lower equilibrium crater density. Similar results were also obtained from the theoretical studies.

The observed population at any one time in Figure 6.23 is, thus, composed of two branches. Small craters follow an equilibrium line with slope $b = 2$. Larger craters follow the steeper production population curve. The inflection point between these two curves is at diameter $D_{eq}(t_1)$ where the production curve crosses the equilibrium line. The right-hand

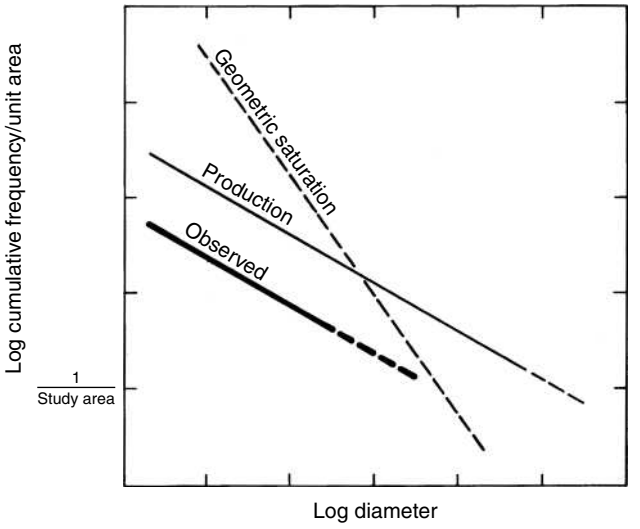

Figure 6.25 Evolution of a crater population with slope $b < 2$. The production population exceeds the geometric saturation line at large diameters. Under these circumstances the evolution of the population is dominated by large impacts that allow a number of small craters to accumulate before wiping the surface clean. The observed population line is always parallel to the production population, but may lie considerably below it. The production population line is dashed at the large-crater end because there must be at least one crater within the study area for the line to be meaningful. The observed population may approach geometric saturation if the entire study area is the site of one large impact, but the population will be dominated by fluctuations because of the statistics of small numbers.

frame in Figure 6.23 shows the crater density at a later time. The crater population is qualitatively similar to that at the earlier time, except that the transition diameter $D_{eq}(t_2)$ is larger. If the rate at which craters accumulate is constant (that is, c in Equation (6.17) is a linear function of time), it is easy to show that the transition diameter $D_{eq}(t)$ grows as $(\text{time})^{1/(b-2)}$. Conversely, if the crater production rate, slope b, and the present D_{eq} are known, the age of the surface can be computed.

6.8.2 $b < 2$ population evolution

Figure 6.25 illustrates the evolution of a crater population with slope b less than 2. Note that in the unlikely event that b exactly equals 2, the observed crater density simply maintains the slope of 2 until the density reaches equilibrium (which occurs simultaneously at all diameters) after which the observed crater density remains constant and the slope remains 2. When b is less than 2 the situation is more complex. In this case the production curve exceeds the geometric saturation line at the large-diameter end of the scale. Gault's (1970) model could not deal with this situation and it took nearly 15 years before the implications of this large-diameter crossing were understood, in spite of the developing knowledge that $b \sim 1.8$ for craters more than about 4 km in diameter on the lunar mare. The situation was

resolved by C. R. Chapman (Chapman and McKinnon, 1986) who performed a Monte Carlo simulation of crater population evolution that included a wider range of diameters than had previously been possible.

Chapman realized that, first of all, the large-diameter end of the cratering curve is dominated by the statistics of small numbers. Even though the production curve apparently crosses the geometric saturation line at all times, enough time must pass after the formation of the initial uncratered surface that at least one large crater has formed on it. The crossing of the two curves makes little sense unless the cumulative number of craters larger than D_s is at least 1 in the finite area under study. Since the probability of a small impact is larger than the probability of a large impact for any $b > 0$, a population of small craters initially develops that follows the production curve closely. Eventually, however, a large impact occurs. With $b < 2$, there is a high probability that this large crater obliterates all or a significant fraction of the study area, wiping out nearly all previous smaller craters. The number of observed craters thus suddenly drops below the production curve. As time passes, small craters again accumulate on the surface. The slope of this new population is equal to that of the production curve but, as shown in Figure 6.25, the cumulative number of such craters is smaller than the production population. The number of small craters continues to grow until the next large impact wipes the slate clean once more.

Under these circumstances there is no "equilibrium" population, although the observed crater density is always well below the geometric saturation limit. The observed crater density fluctuates widely and irregularly, controlled by the large, rare catastrophic-impact events. In spite of these wide variations in density, the slope of the observed population at any given time is roughly equal to that of the production population. The crater densities on a surface of this type are spatially patchy, being low at the sites of recent large impacts and high in areas that have not been struck for a long time by one of the large impacts. Dating such a surface is nearly impossible after the first large impact, unless some area can be found that has escaped all large impacts. In practice, all that can be determined from the crater density on such a surface is the date since the last large impact that affected the particular study region.

6.8.3 Leading/trailing asymmetry

An important assumption in the relative dating of planetary surfaces is that the cratering rate is uniform over the entire body. Although this is true to a high degree of approximation for most bodies in the Solar System, it may be badly violated for synchronously locked planets and moons. Our Moon and the Galilean satellites of Jupiter circulate about their primary with the same face always leading. Just as a car driving into a rainstorm encounters more raindrops on its front windscreen then on the back window, the impact flux on synchronous satellites is higher on the leading side than the trailing side, leading to an asymmetry in cratering rate that may be as large as a factor of 20 or more, depending on the cratering population (Zahnle *et al.*, 2001). The degree of asymmetry depends on the orbits of the impactors. The asymmetry is largest for impactors in orbit about the same primary

or moving at a low relative velocity. In this case the impact velocity on the leading hemisphere is the sum of the orbital velocities of the satellite and impactor, while the velocity on the trailing hemisphere is the difference, so that not only is the flux of impactors different between the leading and trailing hemispheres, but the different impact velocities mean that the same size of impactor yields a different size of final crater. The leading/trailing asymmetry is smaller for heliocentric populations of impactors: Comets, for example, produce much more uniform crater populations than co-orbiting objects on synchronous satellites.

Neptune's moon Triton has a large leading/trailing crater asymmetry, but the Galilean satellites of Jupiter show much less crater asymmetry than expected on theoretical grounds. Some process must, therefore, occasionally reorient their surfaces relative to Jupiter: Either an exchange of their primary-facing and opposite hemispheres due to exchange of their A and B moments of inertia by internal or external (cratering?) mass transfers, or perhaps the momentum impulse of a large cratering event.

6.9 Impact cratering and planetary evolution

Over the last few decades it has become increasingly clear that impact cratering has played a major role in the formation and subsequent history of the planets and their satellites. Aside from their scientific interest, impact craters have also achieved a modest economic importance as it has become recognized that the fabulously rich Sudbury nickel deposit in Ontario, Canada, is a tectonically distorted 140-km diameter impact crater. Similarly, the Vredefort structure in South Africa is a large, old, impact crater. Oil production has been achieved from a number of buried impact craters, such as the 10-km diameter Red Wing Creek crater in the Williston Basin and the 3.2-km diameter Newporte structure in North Dakota. On a more homely level the 60-km diameter ring-shaped depression in the Manicouagan crater in eastern Quebec is currently used as a reservoir supplying water to New York City.

6.9.1 Planetary accretion

Modern theories of planetary origin suggest that the planets and the Sun formed simultaneously 4.6×10^9 yr ago from a dusty, hydrogen-rich nebula. Nebular condensation and hydrodynamic interactions were probably only capable of producing ca. 10-km diameter "planetesimals" that accreted into planetary-scale objects by means of collisions. The timescale for accretion of the inner planets by mutual collisions is currently believed to be between a few tens and one hundred million years. Initially rather gentle, these collisions became more violent as the random velocities of the smaller planetesimals were increased during close approaches to the larger bodies. The mean random velocity of a swarm of planetesimals is comparable to the escape velocity of the largest object, so as the growing planetary embryos reached lunar size, collisions began to occur at several kilometers per second. At such speeds impacts among the smaller objects were disruptive, whereas the larger objects had sufficient gravitational binding energy to accrete most of the material

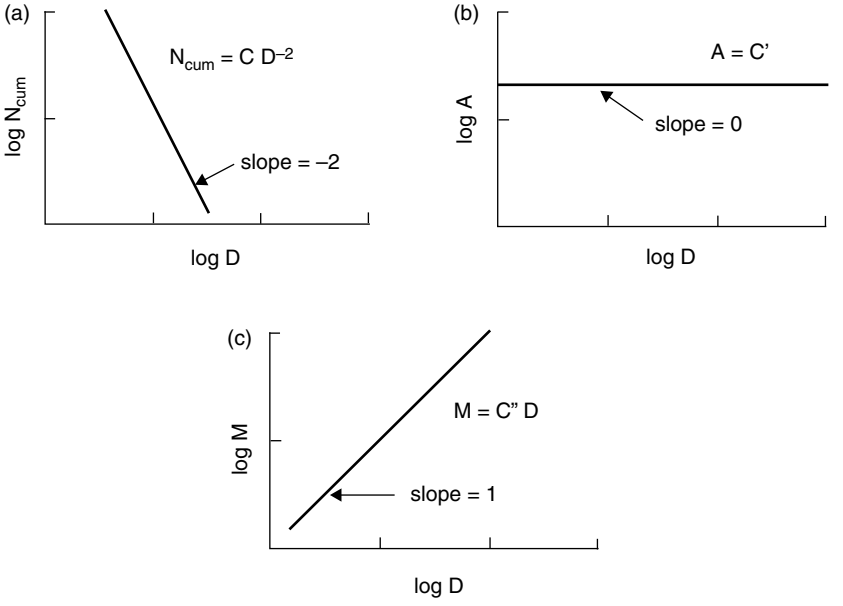

Figure 6.26 The numerical foundation of impact catastrophism. The size-frequency distribution of a typical crater population with slope $b = 2$ is shown in panel (a). In any such distribution there are many more small craters than large ones. Panel (b) shows the area of craters as a function of diameter. This distribution is flat: The total area of small craters and large craters is equal (refer back to Figure 6.20). Panel (c) shows, however, that the mass is concentrated at the large-size scale, so that despite the overwhelming numerical superiority of small craters, the few large ones are the most important in terms of either mass or energy delivered.

that struck them. Infalling planetesimals bring not only mass, but also heat to the growing planets. In the past it was believed that the temperature inside a growing planet increased in a regular way from near zero at the center to large values at the outside, reflecting the increase in collision velocity as the planet became more massive. However, it now seems probable that the size distribution of the planetesimal population was more evenly graded between large and small objects, and that each growing planetary embryo was subjected to many collisions with bodies comparable in size to the embryo itself. Such catastrophic collisions would deposit heat deep within the core of the impacted body, wiping out any regular law for temperature increase with increasing radius and making the thermal evolution of growing planets rather stochastic.

6.9.2 Impact catastrophism

Although small craters (and the impactors that created them) are overwhelmingly more abundant than large ones in terms of numbers, the characteristic $b \approx 2$ slope of most such

distributions in the Solar System implies that most of the mass actually resides at the large end of the size spectrum (see Figure 6.26). This circumstance gives rise to the idea of impact catastrophism: Although small impacts are relatively common, all of the small impacts combined do less geologic work than a few large impacts. This is the formal definition of a catastrophic process and it is a good description of the way that impacts affect the geological evolution of the planets. It also implies that stochastic processes cannot be neglected where impacts play a role, except in situations where b is substantially larger than 2, such as for the Shoemaker–Gault theory of regolith evolution on the Moon.

Note that the usual statistical mechanical idea of energy equipartition does not play a role in the Solar System: The dynamics of the Solar System are such that the time to relax to a state of thermal equilibrium is far, far longer than the age of the Solar System, so that the biggest, most massive objects also carry the most energy.

6.9.3 Origin of the Moon

The origin of the Moon is now attributed to a collision between the proto-Earth and a Mars-size protoplanet near the end of accretion 4.5×10^9 yr ago. This theory has supplanted the three classic theories of lunar origin (capture, fission, and co-accretion) because only the giant impact theory provides a simple explanation for the Moon's chemistry, as revealed in the lunar rocks returned by Apollo. One view of this process is that a grazing collision vaporized a large quantity of the proto-Earth's mantle, along with a comparable quantity of the projectile. While most of the mass of the projectile merged with the Earth (incidentally strongly heating the Earth: If the Earth was not molten before this impact it almost certainly was afterward), one or two lunar masses of vapor condensed into dust in stable Keplerian orbits about the Earth and then later accumulated together to form the Moon.

6.9.4 Late Heavy Bombardment

Sometime after the Moon formed, and before about 3.8×10^9 yr ago, the inner planets and their satellites were subjected to the "Late Heavy Bombardment," an era during which the impact fluxes were orders of magnitude larger than at present. The crater scars of this period are preserved in the lunar highlands and the most ancient terrains of Mars and Mercury. A fit to the lunar crater densities using age data from Apollo samples gives a cumulative crater density through geologic time of:

$$N_{\text{cum}}\,(D > 4\text{ km}) = 2.68 \times 10^{-5}\,[T + 4.57 \times 10^{-7}\,(e^{\lambda T} - 1)] \tag{6.22}$$

where $N_{\text{cum}}(D > 4$ km) is the cumulative crater density (craters/km^2) of craters larger than 4 km diameter, T is the age of the surface in Gyr (T=0 is the present) and $\lambda = 4.53$ Gyr^{-1}. The current cratering rate on the moon is about 2.7×10^{-14} craters with $D > 4$ km/km^2/yr. On the Earth the cratering rate has been estimated to be about 1.8×10^{-15} craters with $D > 22.6$ km/km^2/yr, which is comparable to the lunar flux taking into account the different minimum sizes, since the cumulative number of craters $N_{\text{cum}}(D) \sim D^{-1.8}$. There is

currently much debate about these cratering rates, which might be uncertain by as much as a factor of 2.

Although Equation (6.22) describes an impact flux that decreases monotonically with time, there is currently much debate about the possible reality of an era of "heavy bombardment" between about 4.2 and 3.9 Gyr ago in which the cratering rate reached a local peak (Strom *et al.*, 2005). It is now supposed that eccentric asteroids from the main asteroid belt caused this peak in flux. These asteroids were mobilized as destabilizing orbital resonances swept through the asteroid belt when Jupiter and Saturn underwent an episode of planetary migration. The heavily cratered surfaces of the Moon and terrestrial planets are supposed to have formed at this time. The intense flux obliterated any evidence of an earlier surface on these bodies.

The cratering rates on Mars and Venus are believed to be comparable to that on the Earth and Moon; however, the exact fraction of the Earth/Moon rate is presently uncertain and a subject of controversial discussion. The exact rates will probably remain unknown until radiometric dates on cratered surfaces are determined by means of sample returns from these bodies. Cratering rates in the outer Solar System are even more uncertain and controversial: Most of the craters on the satellites of Jupiter and Saturn are probably formed by comets, whose flux is very uncertain at the small sizes represented by most observed craters (Dones *et al.*, 2009).

The high cratering rates in the past indicate that the ancient Earth must have been heavily scarred by large impacts. Based on the lunar record it is estimated that more than 100 impact craters with diameters greater than 1000 km should have formed on the Earth. Although little evidence of these early craters has yet been found, it is gratifying to note the recent discovery of thick impact ejecta deposits in 3.2 to 3.5 Gyr Archean greenstone belts in both South Africa and Western Australia. Since rocks have recently been found dating back to 4.2 Gyr, well into the era of heavy bombardment, it is to be hoped that more evidence for early large craters will be eventually discovered. Heavy bombardment also seems to have overlapped the origin of life on Earth. It is possible that impacts may have had an influence on the origin of life, although whether they suppressed it by creating global climatic catastrophes (up to evaporation of part or all of the seas by large impacts), or facilitated it by bringing in needed organic precursor molecules, is unclear at present. The relation between impacts and the origin of life is currently an area of vigorous speculation.

6.9.5 Impact-induced volcanism?

The idea that large impacts can induce major volcanic eruptions is one of the recurring themes in the older geologic literature. This idea probably derives from the observation that all of the large impact basins on the Moon's nearside are flooded with basalt. However, radiometric dates on Apollo samples made it clear that the lava infillings of the lunar basins are nearly 1 Gyr younger than the basins themselves. Furthermore, the farside lunar basins generally lack any lava infilling at all. The nearside basins are apparently flooded merely because they were topographic lows in a region of thin crust at the time that mare basalts

were produced in the Moon's upper mantle. Simple estimates of the pressure release caused by stratigraphic uplift beneath large impact craters make it clear that pressure release melting cannot be important in impacts unless the underlying mantle is near the melting point before the impact (Ivanov and Melosh, 2003). Thus, it is probably safe to say that, to date, there is no firm evidence that impacts can induce volcanic activity. Impact craters may create fractures along which pre-existing magma may escape, but themselves are probably not capable of producing much melt. Nevertheless, there were massive igneous intrusions associated with both the Sudbury and Vredefort structures whose genesis is sometimes attributed to the impact, although in this case they may have been triggered by the uplift of hot lower crust. Further study of these issues is needed.

6.9.6 Biological extinctions

The most recent major impact event on Earth was the collision between the Earth and a 10 to 15 km diameter asteroid 65 Myr ago that ended the Cretaceous era and caused the most massive biological extinction in recent geologic history. Evidence for this impact has been gathered from many sites over the last decade, and is now incontrovertible (Schulte *et al.*, 2010). First detected as an enrichment of the siderophile element iridium in the ca. 3 mm thick K/Pg (Cretaceous/Paleogene) boundary layer in Gubbio, Italy, the iridium signature has now been found in more than 100 locations worldwide, in both marine and terrestrial deposits. Accompanying this iridium are other siderophile elements in chondritic ratios, shocked quartz grains, coesite, stishovite and small (100–500 μm) spherules resembling microtektites. All these point to the occurrence of a major impact at the K/Pg boundary. The impact crater is located beneath about 1 km of sedimentary cover on the Yucatan Peninsula of Mexico. Known as the Chicxulub crater, it is about 170 km in diameter and is presently the subject of intensive study.

Further reading

A general survey of all aspects of impact cratering from which several sections of this chapter were abstracted was published by Melosh (1989). This book is presently out of print and often difficult to obtain, but many university libraries possess a copy. A more popular but generally clear and accurate description of terrestrial impact craters can be found in Mark (1987). A good description of the three largest craters on Earth can be found in Grieve and Therriault (2000). Although deeply eroded or otherwise obscured, these craters provide the "ground truth" about impacts that images of extraterrestrial craters cannot give. The surface of our Moon is dominated by impact craters and so the account of its geology by Wilhelms (1987) is necessarily an account of the process of impact cratering. Don Wilhelms also wrote an insightful and very readable historical account of lunar exploration that emphasizes the growing appreciation of impact cratering though the course of the Apollo missions (Wilhelms, 1989). The importance and mechanics of impact crater collapse and the simple-to-complex transition is treated by Melosh and Ivanov (1999), while

the special features of oblique impacts are the subject of a review by Pierazzo and Melosh (2000). The best current summary of the scaling of impact crater ejecta was stimulated by analysis of the ejecta from the Deep Impact Mission cratering event (Richardson *et al.*, 2007). A fine review of the details of crater counting and the application of crater counts to the dating of surfaces appears in a rather unlikely-sounding book, which was the subject of a collaborative project called the *Basaltic Volcanism Study Project* (Project, 1981). The near-field effects of an impact on the Earth are reported by an on-line computer program, whose basic algorithms are described in detail by Collins *et al.* (2005).

Exercises

6.1 Crater dimensions

a) Suppose that the transient crater that collapsed to form the Imbrium Basin on the Moon was D_{tc} = 800 km in diameter. Assuming that it struck vertically at v_i = 22 km/s, what was the diameter of the asteroid that produced Imbrium?

Use the revised Schmidt–Holsapple scaling law (this form implicitly assumes that the target and projectile have the same density):

$$d = 0.671\, D_{tc}^{1.28} \left(\frac{g}{v_i^2} \right)^{0.277} . \qquad \text{(MKS units)}$$

b) If this object struck at 45° to the vertical, what is the period of the Moon's rotation thus imparted, assuming the Moon was free in space (unlikely, but easy to analyze)? Compare this to the Moon's current rotation rate. What can you conclude about the ability of large impacts to unlock the Moon's present synchronous rotation state?

6.2 Crater collapse

a) The empirical relation for the thickness T of an impact crater's ejecta blanket indicates that the average thickness declines as $(R_c/r)^3$, where R_c is the crater radius and

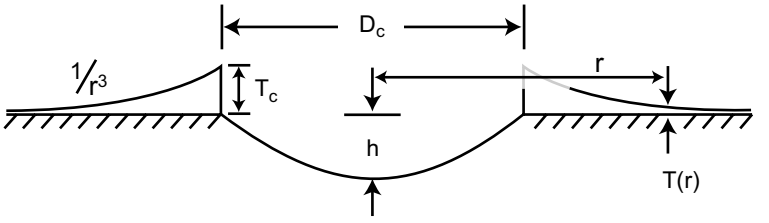

Figure 6.27

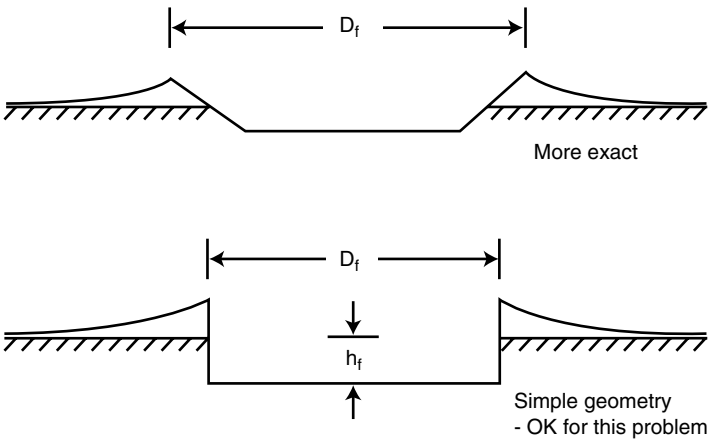

Figure 6.28

r is distance from the crater's center. Assuming that the interior shape of a simple (or transient) crater is described by a paraboloid of revolution with diameter/depth ratio $D_c/h \approx 5$ (use the rim-to-rim diameter D_c and depth h below the original ground surface), and assuming that the volume of the ejecta blanket and crater bowl are equal, derive an equation for the ejecta blanket thickness T_c at the rim. (Hint: It is permissible at this level of accuracy to treat the inner edge of the ejecta blanket as if it were a vertical cliff. An exact result, however, *can* be obtained with a bit of effort, and anyone who can legitimately get the result $T_c=h/5$ should get extra credit for this assignment!)

b) Now suppose the crater collapses to a uniform, constant depth h_f (3 km for the Moon) over its entire interior. Again using volume conservation, how much does the rim-to-rim diameter increase for a 100 km (final diameter) crater?

6.3 Swedish rock rain

In addition to the 5 proven and 33 possible impact crater scars that adorn the country of Sweden, Lilljequist and Henkel (1996) proposed the existence of a truly world-class crater, the Uppland structure, which is supposed to be 320 km in diameter. It encompasses the area around Stockholm up to Uppsala. Using Shoemaker's estimated cratering rate on Earth, N_{cum} ($D > 22.6$ km) = 1.8 x 10^{-15} craters/km^2/yr, and the post-mare size exponent $b = 1.8$ ($N_{cum} \sim D^{-1.8}$), compute the largest crater likely to have formed in Sweden (area ~ 450 000 km^2) since the formation of its surface rocks (1600 to 2300 Myr; use 2000 Myr as a reasonable average).

How does this compare to the size of the putative Uppland structure? Put another way, what is the probability of the formation of a crater as large as Uppland in the past 2000

Myr in Sweden? What does this tell you about the Uppland "crater"? Note that the largest *confirmed* crater in Sweden is the 55 km-diameter Siljan structure.

Using the same cratering flux, compute the maximum size of crater likely to be found in the UK (now it is the student's task to look up the area and mean geologic age of that fair but damp archipelago). Note that a single, probably 4 km diameter, impact crater was discovered in the North Sea by Stewart and Allen (2002) – the Silverpit crater!

6.4 Those blasted Martian rocks!

F. Hörz *et al.* (1999) argued that dish-shaped depressions on large rocks at the Pathfinder landing site on Mars, along with split boulders, might be due to the impact of millimeter to centimeter diameter meteorites at speeds of 300 m/s or more (corresponding to the muzzle velocity of a high-speed rifle).

Evaluate the ability of the Martian atmosphere (surface pressure P_S = 600 Pa) to stop small meteoroids (average pre-atmospheric speed 7 km/s) by comparing the mass of the projectile to the mass of atmospheric gases swept out of the cylindrical volume defined by the meteoroid's path (assume a straight trajectory for the purposes of this estimate).

The total mass of atmospheric gas above a unit area on the surface of a planet is P_S/g, where g is the acceleration of gravity at the surface. Perhaps surprisingly, you do not need to know the density of the atmosphere as a function of height to do this problem (convince yourselves of this!).

Using momentum conservation, estimate how much the meteoroid is slowed upon arriving at the surface as a function of its mass and density. Finally describe how reasonable the Horz–Cintala proposal really is for both stony and iron meteoroids. Could this process be effective on high mountaintops on Earth? Why or why not?

Extra credit

As the meteoroid traverses the atmosphere it encounters a ram pressure of $\rho_a v^2$ on its leading face (ρ_a is the density of the atmosphere, 0.0104 kg/m^3 at the surface of Mars), while its rear is nearly a vacuum for supersonic flight. Estimate the maximum stresses thus acting to crush the meteoroid and compare them to the typical strength of rock or iron for the meteoroids in the problem above. (NB, you *do* have to know the height distribution of density for this problem – you may want to know that the scale height of the Martian atmosphere is about 10 km.)

6.5 Titan gets its kicks!

One of the mysteries about Saturn's large moon Titan is its (slightly) eccentric orbit, e = 0.028. If Titan could dissipate tidal energy as Io or Europa does, its orbit would have circularized long ago. It has been proposed that Titan's Xanadu region is a gigantic impact

structure, 700 km in diameter (with an ejecta blanket bringing the diameter of the entire feature to 1800 km). Using the impact crater scaling relation in Problem 6.1, estimate the size of the object that could have made this crater, assuming a cometary impact velocity of 30 km/s (do not forget the difference between the transient and final crater!). If this object struck the equator of Titan at the optimum place and angle to transform its linear momentum to orbital angular momentum, could this impact have imparted enough momentum to account for Titan's present eccentricity? You may need to know that the orbital angular momentum H of a planet of mass m circling a body of mass M with a semimajor axis of R and eccentricity e is:

$$H = m\sqrt{GMR(1-e^2)}.$$

I will leave you with the task of looking up the values of the necessary parameters. Don't forget that you are looking for a *change* of orbit from circular ($e = 0$) to elliptical. Note that this problem is akin to the oft-asked question of whether the K/Pg impact that killed the dinosaurs could also have knocked the Earth "out of its orbit." What do you think about this possibility?