MERCURY

D. E. Gault, J. A. Burns, and P. Cassen

National Aeronautics and Space Administration, Ames Research Center, Space Science Division, Moffett Field, California 94035

and

R. G. Strom

Department of Planetary Sciences, Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721

INTRODUCTION

Prior to the flight of the Mariner 10 spacecraft, Mercury was the least investigated and most poorly known terrestrial planet (Kuiper 1970, Devine 1972). Observational difficulties caused by its proximity to the Sun as viewed from Earth caused the planet to remain a small, vague disk exhibiting little surface contrast or details, an object for which only three major facts were known: 1. its bulk density is similar to that of Venus and Earth, much greater than that of Mars and the Moon; 2. its surface reflects electromagnetic radiation at all wavelengths in the same manner as the Moon (taking into account differences in their solar distances); and 3. its rotation period is in 2/3 resonance with its orbital period. Images obtained during the flyby by Mariner 10 on 29 March 1974 (and the two subsequent flybys on 21 September 1974 and 16 March 1975) revealed Mercury's surface in detail equivalent to that available for the Moon during the early 1960's from Earth-based telescopic views. Additionally, however, information was obtained on the planet's mass and size, atmospheric composition and density, charged-particle environment, and infrared thermal radiation from the surface, and most significantly of all, the existence of a planetary magnetic field that is probably intrinsic to Mercury was established.

In the following, this new information is summarized together with results from theoretical studies and ground-based observations. In the quantum jumps of knowledge that have been characteristic of “space-age” exploration, the previously obscure body of Mercury has suddenly come into sharp focus. It is very likely a differentiated body, probably contains a large Earth-like iron-rich core, and displays a surface remarkably similar to that of the Moon, which suggests a similar evolutionary history.

1 Permanent Address: Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853
SIZE AND MASS

Mariner 10's close flybys of Mercury have permitted the planet's mass to be ascertained accurately. Esposito et al. (1976) have deduced that the ratio of the solar mass to Mercury's is $6,023,600 \pm 600 \left[3.3020 \pm 0.0037 \times 10^{26} \text{ g}\right]$ from the first flyby and $6,023,700 \pm 300$ from the third (cf Howard et al. 1974); the second encounter was at too great a distance from Mercury to provide a useful estimate of its mass. These values lie near the much larger bounds (5,972,000 ± 45,000) of previous results (cf Duncombe, Klepczynski & Seidelmann 1973, Duncombe, Seidelmann & Klepczynski 1973), the most accurate of which came from a simultaneous solution of optical and radar positions for the inner planets (Ash, Shapiro & Smith 1971).

The spacecraft orbit also defined $J_2$, the oblateness parameter, for the first time. Based on the third encounter, $J_2$ equals $(8.6 \pm 6) \times 10^{-5}$, whereas the orbit from the first flyby, which was less suited for analysis, gives a value of the same order (Esposito et al. 1976). For most other planets, $J_2$ is closely associated with rotational flattening and gives information on the central condensation of the planet. Because Mercury's spin is so slow, however, the nonhydrostatic part of $J_2$ is larger than the centrifugal part ($\sim 10^{-6}$) by almost two orders of magnitude, and, unhappily, little can be learned of the interior from the value of $J_2$ except that its strength is comparable to that of other terrestrial planets. The current data show structure in the local gravity field and P. B. Esposito (personal communication 1976) believes that the data set could perhaps generate a very rough value of $J_4$ with considerably more effort.

Because of its smallness and its angular proximity to the Sun (its maximum elongation being 27°), Mercury is a difficult object to study telescopically, with its disk merely 6.9" to 10.9" across. Thus classical estimates of its radius are necessarily imprecise. Nevertheless, they have been surprisingly correct, giving estimates of $2440 \pm 7.5$ km (de Vaucouleurs 1964). Refined radar ranging experiments have given an accurate mean radius of $2439 \pm 1$ km (Ash, Shapiro & Smith 1967, 1971), which has been confirmed by the values determined from the occultation by Mercury's disk of Mariner 10 dual frequency radio signals during first encounter: $2439.5 \pm 1$ km at 1.1° N, 64.7° E (nightside) and $2439 \pm 1$ at 67.6° N, 258.4° E (dayside) (Howard et al. 1974, Fjeldbo et al. 1976).

Accepting the Mariner 10 value for the mass and the radar measurement of the radius produces a planet with a mean density of $5.433 \pm 0.012$ g cm$^{-3}$, a value essentially equal to that of the Earth but corresponding to the highest known uncompressed density for a solar system body.

ORBIT

At present Mercury's orbit is distinguished from those of the other planets in the smallness of its semimajor axis ($a = 0.387$ au) and by its comparatively large eccentricity ($e = 0.206$) and orbital inclination relative to the ecliptic ($i = 7.0°$). With
the exception of the distant Pluto. Mercury's orbit is the most inclined and elliptic
of any planet. This is not a temporary condition; the secular theory of Brouwer &
van Woerkom (1950), which accounts for the mutual perturbations of the planets
(excluding Pluto), has been numerically integrated by Cohen, Hubbard &
Oesterwinter (1973) and gives a mean $e = 0.175$ and $i = 7.2^\circ$ over $10^7$ years, values
that are about three times larger than for other planets. The half amplitude of the
computed periodic oscillations in $e$ is about 0.07 and has a prominent period of about
$10^6$ years, while $i$ varies $\pm 2^\circ$ with the largest period at nearly $10^6$ years and much
smaller amplitude variations ($\pm 0.2^\circ$) occurring in a period of $10^5$ years.

The conventional wisdom is that the unusually large $e$ and $i$ are mere chance
occurrences or, at the most, that the mercurian inclination shows a primordial tie
to the solar equatorial plane, which is tilted at $6^\circ$ to the invariable plane, and,
perhaps, that the high $e$ is caused by the high-collision velocities between proto-
Mercury and planetesimals perturbed into the inner solar system by Jupiter (Kaula
1976).

Ward, Colombo & Franklin (1976), however, give another explanation for the
odd $e$ and $i$. They demonstrate that a secular resonance between the precession
rates of the lines of apsides for Mercury and Venus would have existed in the past
if and when the Sun had an oblateness $J_2 \sim 10^{-2}$, corresponding to a spin period
of about five hours; nearly the same $J_2$ would produce another resonance between
the precession rates of the lines of nodes of the two planets. During such secular
resonances, Mercury's $e$ and $i$ are pumped up by Venus. If but only if -the
characteristic time scales for decay of the solar rotation is about $10^6$ years is it
possible for this mechanism to produce the presently observed $e$ and $i$. From a
comparison with T-Tauri and other young stars, such a rapid solar rotation and
decay rate are not implausible for the early Sun. Van Flandern & Harrington (1976)
boldly suggest that Mercury may be an escaped former satellite of Venus, but they
are unable to find a mechanism to account for the large difference between their
semimajor axes.

The prograde precession of Mercury's perihelion by $43^\circ$ more per century than
can be explained by the Newtonian action of the other planets was one of the major
failings of classical celestial mechanics. Einstein's general theory of relativity
accounted for this discrepancy and its precise description of the motion of the inner
planets still remains one of the most stringent tests to be passed by any challenger
to Einstein's theory (Will 1974).

**ROTATION**

For nearly a century following Schiaparelli's claim in 1889 that Mercury rotated
slowly, telescope observers agreed that the rotation period was synchronous with
the orbital period of 88 days (see Colombo & Shapiro 1966, Smith & Reese 1968,
and historical references therein). This result was overturned in 1965 by the radar
measurement of a rotation period of $59 \pm 5$ days by Pettengill & Dyce (1965),
subsequently refined to $58.65 \pm 0.25$ days (Goldstein 1971). After theoretical justi-
fication that solid-body tides due to the Sun's attraction might well slow the spin
to a value somewhat higher than the synchronous rate because of the high orbital eccentricity (Peale & Gold 1965), optical astronomers—in one of the better recoveries ever recorded—reanalyzed historical observations to find a spin period of 58.644 ± 0.009 days (Murray, Dollfus & Smith 1972). These error bounds enclose the 58.6457-day period corresponding to a 2:3 commensurability between the axial and orbital period. Colombo (1965) had suggested that such a commensurability would result from the combined effect of tides and solar torques on the nonspherical shape of the planet, and Liu & O'Keefe (1965), Colombo & Shapiro (1966), and Goldreich & Peale (1966) showed it to be a stable motion. Klaasen (1975, 1976) obtained 58.6461 ± 0.005 days from shadows measured on consecutive Mariner 10 encounters to confirm the resonance.

Ground-based telescopes and radar and spacecraft observations have also been used to find the planet's obliquity, the angle between the rotation axis and pole of the orbit. Telescopic work generally suggests that the rotation axis is perpendicular to the orbit plane to within 3° (Murray, Dollfus & Smith 1972), in agreement with much cruder radar results (Dyce, Pettengill & Shapiro 1967). Analysis of Mariner 10 pictures (Klaasen 1976) also gives a nearly perpendicular rotation axis (obliquity = 2° with a 2.6° × 6.5° error ellipse). Although a very precise determination (<1') of the obliquity would permit an evaluation of \( (C-A)/C \), where \( A < B < C \) are the planet's principal moments of inertia, this precision cannot be achieved without the use of a lander (Peale 1972). Besides accurately determining the obliquity, a lander could measure the amplitude of the libration in longitude which, along with accurate values for the gravitational coefficients \( J_2 \) and \( C_{22} \), may be able to indicate the extent of a molten core (Peale 1976b).

Presumably, Mercury did not originate in this spin state. Most likely it had been rotating much faster, perhaps with a period near eight hours, like so many solar system bodies today (cf Burns 1975), but it has been subsequently slowed by solar tides (Burns 1976). Such tides despin Mercury with a characteristic decay time of a billion years for a tidal dissipation factor \( Q \) of 30. This slowing would have heated the interior by about 100° K and would have caused substantial surface strains (corresponding to a change in radius of tens of kilometers) that, if not relaxed, would result in surface stresses well above fracture limits. Surface features characteristic of this failure mode (normal faulting) are not seen today although a global structural pattern appears to be present as linear features (D. Dzurisin, personal communication). Large extensional surface strains, equivalent to an increase in radius of 13 km, are generated during core infall while smaller compressional strains (\( \Delta R \approx 2 \) km) subsequent to core formation accompany the planet's cooling; the latter may account for the common lobate scarps (Solomon 1976). Other important volume changes, largely unexamined, could occur because of possible changes in phase of the core.

The fact that Mercury spins slowly may explain the absence of natural satellites about it. Both Burns (1973) and Ward & Reid (1973) have pointed out that tides on a slowly spinning planet like Mercury impose drag forces on satellites that decay their orbits, causing the satellites to impact eventually onto the planet.

The most complete treatment of the 2:3 resonant rotation (Goldreich & Peale...
1968 and earlier papers) demonstrates that such a rotation is stable as long as the solar torque exerted on the planet's asymmetric equatorial shape is always larger than the time averaged tidal torque, or \((B-A)/C > 10^{-7} Q^{-1}\). Since \(J_2 = \frac{C-\frac{1}{2}(B+A)}{MR^2}\) is likely to be comparable to \((B-A)/C\), the Mariner 10 result of \(J_2 \sim 10^{-4}\) assures that the stability criterion is almost certainly satisfied. The \(r^{-3}\) dependence of the torque on the permanent deformation, together with the eccentric orbit, aligns the long equatorial axis with the Sun-Mercury line at perihelion. Since many other commensurate spin rates are also stable for reasonable choices of \(Q\), the question arises as to why Mercury avoided capture into these as it was tidally slowed. Goldreich & Peale (1966) show that capture occurs when the rotational energy lost (by dissipation) over one libration—whether due to tidal or internal origin—exceeds the kinetic energy stored in the angular motion measured relative to that commensurate spin state. Assuming the rotational phase to be randomly distributed, the capture probability into a particular state will be the ratio of these two energies and will therefore depend on the energy loss mechanism. A comforting confirmation of the theory is the fact that the capture probability of Mercury into a 2:3 resonance is much greater than into any higher-order resonance for the most common tidal models. The putative molten core would dissipate further energy through shear losses at the core boundary. Peale & Boss (1977) point out that the escape of Mercury from capture into the 1:2 resonance requires that, if a molten core does exist, its kinematic viscosity must be comparable to that of water \((0.01 \text{ cm}^{-2} \text{ sec}^{-1})\) and \(Q \lesssim 100\).

Colombo (1966) and Peale (1969) also have shown that a generalization of Cassini's laws applies to Mercury: There are three possible obliquities for which the spin axis lies stably in the plane formed by the orbit normal and the axis about which the orbit precesses (i.e. the invariable plane of the solar system). Peale (1974, 1976a) deduces from a study of the history of the obliquity that all initial conditions (except for a pathological case of slow primordial spin, large initial obliquity, and, effective core-mantle interaction) produce a final state where the spin axis lies near the normal to the orbit as observed.

**ATMOSPHERE**

The atmosphere is very tenuous; the surface pressure at the subsolar point is less than a few times \(10^{10}\) mb. Such an atmosphere is essentially exospheric—the gas is expected to be collisionless right down to the planetary surface.

Helium and atomic hydrogen have been identified as constituents of Mercury's atmosphere by detection of their emission lines by the UV spectrometer aboard Mariner 10 (Broadfoot et al. 1974, 1976, Broadfoot 1976). The absence of other UV emission lines and the lack of measurable absorption during solar occultation have established upper limits for several other possible constituents: \(\text{H}_2\), \(\text{O}_2\), Ne, Ar, O, \(\text{CO}_2\), \(\text{H}_2\text{O}\), and \(\text{N}_2\) (see Kumar 1976 for a useful comparison with the lunar atmosphere). In addition, an upper limit to the electron density of \(10^3\) cm\(^{-3}\) has been determined by the Mariner 10 radio occultation experiment (Fjeldbo et al. 1976).
A model helium atmosphere has been calculated by Hartle et al. (1975) (see also Hartle et al. 1973, Hodges 1974), in which neutral particles follow ballistic trajectories and possess a Maxwellian distribution of energies characterized by the local surface temperature. It is assumed that the surface is saturated in helium and therefore cannot act as a sink. Because the nightside surface is about 100°C, 450°C colder than the dayside, particles spend more time on the nightside and cause the density to be higher there. Hartle et al. find a night-to-day surface density ratio of 200. However, this ratio is apparently too large, and the predicted nightside scale height is too small to match the observations of Mariner 10, which suggests that the surface boundary condition may be more complex than assumed (Broadfoot et al. 1976).

The solar wind is a likely source of helium, for, although most of it is excluded from the magnetosphere, less than a tenth of a percent of its flux need be neutralized at the surface in order to fulfill the source requirement, if the dominant loss mechanism is thermal escape. It seems likely that this small amount could enter the magnetosphere via the tail, through the cusp regions, or by diffusion across the magnetopause. On the other hand, the radioactive decay of uranium and thorium in the crust may also be sufficient to maintain the low helium density.

The hydrogen component is interesting in that two populations appear to be present: "thermal" gas, which has an exospheric structure near that expected from equilibrium with the surface, and "nonthermal" gas, which has a much smaller scale height than the former, and predominates at altitudes below 300 km. The reactive nature of hydrogen makes the problem of identifying its sources and sinks difficult: Thomas (1974) has suggested that photolysis of water may play an important role in the hydrogen budget. Model atmospheres that include such a process have yet to be constructed.

It is possible that photo ionization followed by migration to the magnetopause, where the ions are swept away in the solar wind, is an important loss mechanism. This is particularly true for heavy gases such as neon, argon, carbon dioxide, and water, whose thermal escape times are long. Kumar (1976) has estimated upper limits to their supply rates imposed by the observed density limits, and the assumption that photo ionization followed by solar wind loss is the primary sink. The problem is complicated by the possibilities that these gases can condense on the nightside or collect in permanently shaded areas near the poles, and that surface interactions may not be in steady state. Nevertheless, in the cases of CO$_2$ and H$_2$O at least, the supply rates appear to be much lower than would be expected if Mercury were outgassing at a rate comparable to the Earth. Either Mercury is intrinsically volatile deficient (as would be consistent with Lewis' 1972 compositional model), or the planet is merely inactive, at least in its outer layers.

**MAGNETIC FIELD AND MAGNETOSPHERE**

The magnetometer aboard Mariner 10 revealed the existence of a weak but apparently permanent magnetic field associated with Mercury (Ness et al. 1974, 1975, 1976). Although the strength of the field is much less than that of the geomagnetic field, it is strong enough to form a magnetosphere similar to Earth's in
several respects. Bow-shock, magnetosheath, and magnetopause are identifiable in
the magnetic and plasma data (Ness et al. 1974, 1976, Ogilvie et al. 1974); the near-
planet field is well represented by a dipole (with moment $5 \times 10^{22}$ G cm$^3$) aligned
approximately parallel to the rotation axis, and with the same polarity as the Earth's
field. There are indications that the magnetosphere is stretched in the antisolar
direction, possibly forming a long magnetotail with an imbedded neutral sheet
(although Mariner 10 did not explore this region). Furthermore, there is evidence
that Mercury's magnetosphere, like Earth's, is subject to substorms--complex
transient distortions caused by interaction with the solar wind (Siscoe et al. 1975).

There is, however, an important difference between Mercury's and Earth's
magnetospheres. Mercury occupies much more of its magnetosphere than does the
Earth. The average distance from the center of Mercury to the nose of the magnetos-
sphere is only 1.6 planetary radii, whereas for Earth this distance is more than 10
planetary radii. Therefore, charged particles within the magnetosphere encounter
Mercury much more readily than their counterparts intercept the Earth, with the
result that there is no trapped radiation region similar to the Van Allen belts. There
may in fact be infrequent periods when the solar wind is strong enough to compress
the field to the surface of the planet (Siscoe & Christopher 1975).

The source of Mercury's magnetic field has not been established. An active
magnetohydrodynamic dynamo within the planet is a favored hypothesis, even if
only by analogy with other large-scale sources of magnetism in the solar system--
the Sun, Earth, and Jupiter. However, neither theory nor observations have yet
provided a means of predicting the magnetic field produced by a fluid dynamo
using a specified energy source: only the most rudimentary requirements can be
defined at this time. It is certain that the magnetic Reynolds number $R = 4 \pi \sigma L/Vc^2$
must exceed unity in order that fluid motions can build up magnetic field faster
than it diffuses. (The electrical conductivity is denoted by $\sigma$, $c$ is the speed of light,
and $L$ and $V$ are typical length and velocity scales.) Applied to planetary interiors,
this requirement immediately implies molten material of high electrical conductivity.

For Mercury, iron is a sufficiently good conductor; molten silicates are probably
not (Stevenson 1975). Rotation has long been cited--but never proved to be--a
necessary condition for astrophysical dynamos. The amount necessary may be
minimal: only a slight amount of angular velocity of a planetary-sized dynamo
(of low viscosity) is necessary for the Coriolis force to dominate the momentum
balance, the condition that properly characterizes "rapid rotation" (e.g., see Gubbins
1977). Interestingly, this theoretical result has emerged only after the discovery of
the mercurian field.

The power necessary to maintain a dynamo must be greater than that dissipated
within it. For Mercury, this is probably about a few times $10^{10}$ ergs sec$^{-1}$ (Gubbins
1977), which is not severe. However, the energy necessary to maintain a molten
core is substantial, and is discussed later.

There are other possible sources for the mercurian field besides a dynamo
(Stevenson 1974). For instance, the field may be due to the permanent magnetization
of iron-bearing rocks in the outer layers of the planet. Such magnetization is found
in rocks from the Earth, Moon, and meteorites, and is most commonly acquired by
the cooling of the rock through the Curie temperature in the presence of an ambient field. (The Curie temperature of iron is 770°C.) If there was once a dipolar field due to a dynamo that has since died out (perhaps by depletion of its energy source of solidification of the core), a dipolar remanent field would remain, its source being the residual moment retained by those layers of the planet that cooled through the Curie point while the dynamo was active. Stephenson (1976) has determined the degree of magnetization that would be necessary for such a field to be comparable to that measured by Mariner 10 and concludes that the mercurian field might indeed be a remanent one.

The interplanetary electromagnetic field can generate measurable planetary fields by induction, as occurs in the Moon (Sonett & Colburn 1968, Colburn, Sonett & Schwartz 1972). However, such an induction mechanism is incapable of providing as strong and steady a field as that observed by Mariner 10 (Herbert et al. 1976, Ness et al. 1976).

SURFACE FEATURES

Mariner 10 acquired over 2700 useful pictures of Mercury during its three encounters, which covered approximately 50° of the surface at resolutions varying from about 100 meters to 5 km. Although this portion of Mercury's surface is remarkably similar to the Moon's, there are significant differences that indicate important departures between the surface evolution of Mercury and the other terrestrial planets.

In general, the surface of Mercury is pockmarked with craters ranging in diameter from at least 100 m (highest resolution) up to basins 1000 km across (Figure 1). They present a spectrum of degradational types similar to those on the Moon, and many exhibit extensive ray systems. Large areas of lightly cratered smooth plains similar to the lunar maria fill and surround the major basins. However, the albedo of this material is similar to the more heavily cratered terrain, and therefore the contrast is considerably less than that between the lunar maria and highlands. Furthermore, mercurian crater morphology is somewhat different from that of lunar craters, probably as the result of Mercury's greater surface gravity (Gault et al. 1975). Large lobate scarps occur in most areas of Mercury and are probably of tectonic origin. These scarps are unique to Mercury and indicate that the tectonic style and history has been quite different from that of the other terrestrial planets (Strom, Trask & Guest 1975).

Major Physiographic Provinces

Mercury's surface can be divided into three major physiographic provinces: intercrater plains, heavily cratered terrain, and smooth plains. In addition to these major terrain types, several localized units can be identified. By far the most interesting is the hilly and lineated terrain which is antipodal to the large, relatively fresh Caloris basin. Figure 2 is a generalized terrain map of Mercury modified from Trask & Guest (1975) which shows the distribution of these terrain types.

The intercrater plains are probably the most widespread terrain unit on Mercury. They occur between and around clusters of large craters comprising the heavily
MERCURY 105

cratered terrain, and are characterized by a level to gently rolling surface covered by a high density of superposed small craters in the size range 5–10 km (Figure 3a). Many of these small craters form chains or clusters, and individual craters are often elongated and open at one end. These characteristics are common to secondary impact craters and therefore it is probable that most of these superposed craters are of secondary origin. The age of the intercrater plains relative to the other units is of paramount importance for deciphering the surface history of Mercury. Trask & Guest (1975) consider the intercrater plains to be older than the heavily cratered terrain and therefore the oldest exposed unit on Mercury because 1. they do not appear to embay or cut individual craters of the heavily cratered terrain, and 2. the only apparent source of the many secondary impact craters on the intercrater plains appears to be the large craters and basins of the heavily cratered terrain. However, at several locations within the intercrater plains are isolated, partially buried craters similar to the “ghost rings” found in the lunar maria. This suggests that at least in part the intercrater plains may be younger than the heavily cratered terrain. More detailed studies of this terrain type are required before firm conclusions can be reached regarding its relative age.

The origin of the intercrater plains is even more uncertain than their age. A similar terrain of very restricted distribution occurs on the Moon southwest of the Nectaris basin and is called the pre-Imbrium pitted plains (Wilhelms & McCauley 1971). This terrain has been interpreted as basin ejecta by Howard, Wilhelms & Scott (1974), and a similar interpretation has been applied to the mercurian intercrater plains by Wilhelms (1976). However, a major difference between the Moon and Mercury is that the mercurian intercrater plains are the most widespread terrain on Mercury (see Figure 2) and cover a much greater area than their lunar equivalent. Furthermore, Mercury appears to have fewer multi-ring basins in the size range 200–1000 km (Murray et al. 1974) than the Moon, and the ejecta from these basins has a more restricted ballistic range than on the Moon (Gault et al. 1975) because of Mercury's greater surface gravity. This great areal distribution of intercrater plains, restricted ballistic range, and apparent paucity of source basins argue against a basin ejecta origin for the mercurian intercrater plains. If the majority of intercrater plains are the oldest unit and predate the heavily cratered terrain, then they may represent an ancient, primordial surface that does not have a lunar counterpart because of the greater density of lunar basins and greater ballistic range of ejected lunar material (Trask & Guest 1975). An alternative explanation is that if much of the intercrater plains are younger than the heavily cratered terrain, they may represent a volcanic episode intermediate in age between the formation of the heavily cratered terrain and smooth plains. A complicating factor in choosing between interpretations is the unknown degradational effects of the global-wide seismic disturbances produced as a result of the monstrous impact event that formed the 1300-km diameter Caloris basin (Guest & Gault 1976). Clearly more detailed investigations are required to decide between competing hypotheses.

The heavily cratered terrain primarily consists of clusters of closely packed, overlapping craters ranging from about thirty to several hundred kilometers in diameter (Figure 3a). Ejecta blankets and discrete secondary crater fields are
generally lacking around craters comprising this terrain. The interiors of many of the craters are filled with smooth plains that are much less cratered and younger than the adjacent intercrater plains. In appearance the terrain is very similar to the lunar highlands and undoubtedly resulted from a heavy bombardment of small planetesimals early in Mercury’s history.

The mercurian smooth plains form sparsely cratered, relatively level surfaces

Figure 1. Photo mosaics of high-resolution pictures of Mercury obtained during Mariner 10’s first encounter with the planet on 29 March 1974. (a) “Incoming” hemisphere visible as spacecraft approached the planet. The equator lies about 20° above the center of the photo, and the evening terminator lies near 10° west longitude. The surface exhibits a heavily cratered terrain totally unlike the “outgoing” hemisphere. NASA No. 74-H-239 JPL Photo. (b) “Outgoing” hemisphere visible after spacecraft passed planet. The equator lies about 20° below center of the photo with the morning terminator near 190° west longitude. The 1300-km diameter Caloris basin and large areas of smooth plains contrast with the older heavily cratered terrain in the incoming hemisphere. NASA No. 74-H-253 JPL Photo.
(Figure 3b), which in some respects are similar to the maria (Strom, Trask & Guest 1975) and in other respects are similar to the lunar light plains known as the Cayley formation (Wilhelms 1976). Overlap relations and the density of superimposed craters indicate that they are younger than the more densely cratered areas and among the youngest surfaces on Mercury (Trask & Guest 1975, Guest & Gault 1976). Craters in the size range 5–10 km are much less abundant than on the intercrater plains. Numerous ridges, which grossly resemble both lunar and martian wrinkle ridges, occur in the large expanses of smooth plains. However, the mercurian ridges generally lack the crenulated crests so characteristic of lunar and martian ridges. The most extensive tracks of smooth plains occur in and around the Caloris and north polar basins (Figures 1b and 2b). Other patches mostly occupy the floors of basins and large craters. Unlike the lunar maria, the albedo of the smooth plains does not contrast sharply with the surrounding heavily cratered terrain or intercrater plains (Hapke et al. 1975).

![Figure 2](image-url)  
Figure 2 Generalized geological terrain map of Mercury modified from Plate 1 of Trask & Guest (1975). (a) Incoming hemisphere; (b) Outgoing hemisphere.
Two possible origins have been proposed for the smooth plains: volcanic and basin ejecta. Evidence cited in support of a volcanic origin includes: 1. the large volume of smooth plains; 2. the great differences in volume of the plains material around basins of comparable size; 3. the similarity in morphology and distribution between the smooth plains and the lunar maria; 4. stratigraphic relationships which indicate that the smooth plains are younger than many of the basins in which they lie; 5. the large volume of plains peripheral to the north polar basin, making derivation by impact from that basin unlikely; 6. contrasting albedo and

Figure 3  Examples of the major physiographic provinces. (a) Intercrater plains (IC) and heavily cratered terrain (HC) centered at 56°S, 128°W. View is about 400 km across. After Trask & Guest (1975). (b) High-resolution view about 320 km across showing smooth plains (SP) and hummocky plains (HP) about 500 km east of the Caloris basin. Part of a pre-Caloris basin 240 km in diameter in the lower part of the scene is filled with smooth plains. At the western margin of this basin is a west-facing scarp that forms the boundary between hummocky plains (west) and smooth plains (east). After Strom, Trask & Guest (1975). (c) High-resolution view of the hilly and lineated terrain that is antipodal to the Caloris basin. Scene is about 500 km across centered at 31°S, 19°W. (d) Photomosaic of one of the most prominent lobate scarps (Discovery Scarp). This feature is about 550 km long and transects two craters 55 and 35 km in diameter. Maximum height of scarp is about 3 km.
color of some of the plains material compared to the immediate surroundings; and 7. a lack of source basins for plains filling older basins (Strom, Trask & Guest 1975, Trask & Strom 1976). Wilhelms (1976), on the other hand, considers that the smooth plains more closely resemble the lunar light plains, which are probably largely basin ejecta. He cites similarities in stratigraphic relations, surface morphology, and albedo contrasts between the smooth plains and lunar light plains (Cayley formation) as evidence of a basin ejecta origin for the mercurian deposits, and suggests that the ejecta behaved more like a fluid than did lunar basin ejecta because of the greater surface gravity of Mercury. The question of mercurian volcanism should be kept open pending more conclusive observational evidence or more convincing theoretical arguments. However, the evidence for volcanism outlined above, together with thermal history models required to account for at least a partially molten core apparently necessary to explain the magnetic field, makes past mercurian volcanism a good working hypothesis.

A peculiar topography of localized distribution comprises one of the most unusual terrains on Mercury. This peculiar set of landforms occurs only in one area viewed by Mariner 10, which is antipodal to the Caloris basin. It consists of hills 5-10 km wide and 0.1-1.8 km high and several large linear valleys arranged in an orthogonal pattern (Figure 3c). The terrain includes crater rims that have been broken up into hills and depressions. However, the floors of many of these craters are occupied by smooth plains that have not been affected by the process that caused the hilly and lineated terrain. Because somewhat similar terrain occurs at the antipodal regions of the Imbrium and Orientale basins on the Moon, Schultz & Gault (1975) have postulated that the terrain may have formed at the same time as the Caloris basin by the focusing of seismic energy at the antipodal point. Moore et al (1974) have attributed the lunar terrain to clustering of basin-related secondary impacts at the antipode, but this is an unlikely process on Mercury due to the requisite ejection velocities (3.5-4 km sec⁻¹) necessary for ballistic transport to the antipode. Alternatively, Wilhelms (1976) has suggested that the intersecting linear structures are secondary crater chains formed by ejecta from mercurian basins hidden in the terminator.

**Structure of the Caloris Basin**

The Caloris basin is the largest structure viewed by Mariner 10 (Figure 4). It is 1300 km in diameter and resembles the lunar Imbrium basin in both size and morphology (Strom, Trask & Guest 1975). The interior of the basin is occupied by smooth plains that are highly fractured and ridged. The basin perimeter is defined by a ring of irregular mountains averaging about 2 km in height above the floor. A weak outer scarp occurs at a distance of 100-160 km from the main scarp in the northeastern part of the basin, and between these two scarps lie relatively smooth hills and domes. An extensive system of valleys and ridges radiates from the basin for a distance of about one basin diameter and strongly resembles the lunar Imbrium radial system. Both radial valleys and old craters are embayed by plains material that surrounds the basin out to three basin radii. This plains material can be divided into two units: smooth plains and hummocky plains (Trask &
The texture of the hummocky plains and their proximity to the basin rim (Figure 3b) suggest that they are basin ejecta excavated from the basin by the impact that formed it. As discussed previously, the smooth plains may be in part or whole volcanic flows emplaced after the Caloris event (Strom, Trask & Guest 1975, Trask & Strom 1976), or a smooth ejecta facies similar to the lunar light plains (Wilhelms 1976).

The floor structure of the Caloris basin appears to be unique; no basin on the Moon or Mars shows this type of structure. It is characterized by a high density of fractures and ridges that both show two primary orientations: one concentric with respect to the border of the basin, and the other radial. Furthermore, the intensity of the fracturing progressively increases basinward. Both fractures and ridges are probably due to readjustment of the basin floor subsequent to emplacement of the floor material.

*Figure 4*  Photomosaic of the Caloris basin showing the highly ridged and fractured floor of the interior, the well-developed radial system of scarps in the northeastern portion of the basin, and the extensive areas of smooth plains surrounding the basin.
Tectonic Framework

The tectonic framework of Mercury is characterized by large and widely distributed lobate scarps (Strom, Trask & Guest 1975). These features are unique to Mercury and form relatively steep escarpments that show a broadly lobate outline on a scale of a few to tens of kilometers (Figure 3d). They vary in length from about twenty to over five hundred kilometers and have heights of a few hundred meters to about three kilometers. The crests of the scarps are rounded in contrast to the sharp crests formed by normal faulting and graben on the Moon and Mars. They often transect a variety of terrain types, and in at least one case a crater rim has been offset by about 10 km and appears to have had its radius shortened. These transectional and morphological relationships indicate that the majority of the scarps are probably reverse or thrust faults due to compressive stresses. However, about 17% of the scarps are confined to smooth plains on crater floors or occur at the interior margins of old basins. The origin of these scarps is not clear; they could be flow fronts rather than fault scarps.

There are only two local areas on Mercury that show evidence of tensional stresses. One of these areas is the highly fractured floor of the Caloris basin, while the other is the hilly and lineated terrain antipodal to it. Therefore, the surface of Mercury is dominated by structures indicative of compressive stresses.

With the exception of the small percentage of scarps confined to crater floors, lobate scarps of probable tectonic origin appear to have a rather uniform distribution over that part of Mercury viewed by Mariner 10. If this region of Mercury is representative of the planet as a whole, then lobate scarps probably have a global distribution. Therefore, the entire planet appears to have been subjected to compressive stresses resulting in a general crustal shortening. Estimates of the amount of horizontal displacement represented by lobate scarps suggest that there has been a decrease in surface area of about $6.3 \times 10^4$ to $1.3 \times 10^5$ km$^2$, which represents a decrease in radius of about 1.2 km (Strom, Trask & Guest 1975). This value is consistent with a 2-km decrease in radius based on thermal history models, which predict a contraction of the lithosphere due to cooling after core formation (Solomon 1976).

Craters

The morphology of the mercurian craters, down to the limits of the photographic resolution, is grossly similar to that of their lunar counterparts. The craters unquestionably represent, as on the Moon, the end products of impacts by the complete size spectrum of meteoritic material, and differences between the two families of craters are attributed by Gault et al. (1975) to the factor of 2.2 difference in the respective gravitational acceleration environments. The primary morphologic elements (rim ejecta deposits, fields of satellitic secondary craters, terraces on inner walls, central peak(s) and/or inner rings of mountains) that characterize lunar craters are all associated with craters on Mercury. Moreover, the morphology changes from sharp to soft with increasing age of the features, as on the Moon, and the crater elements ultimately disappear or become barely recognizable as the crater is
degraded to a low rim that is frequently partially obliterated by superposed craters. Thus, the primary degradational process on Mercury is probably the same as on the Moon; i.e. erosion and ballistic sedimentation caused by meteoritic bombardment.

The smallest craters are predominantly bowl-shaped, and with increasing size the craters exhibit a systematic progression of morphologic types beginning with the appearance of central peaks and terracing on the inner walls (Figure 5a and 5b). Further increases in size are accompanied by the appearance of complex structures of central peaks which undergo a transition into an inner ring of mountains that is

Figure 5  Examples of the morphologic progression of mercurian craters with increasing size. (a) Small bowl-shaped craters to larger features with incipient terraces and central peak(s). The fresh, sharp crater (centered) is a 20 km diameter structure with Mercury's surface longitude reference crater Hun Kal on its southern flank indicated by arrow. (b) A 98-km diameter crater illustrates the typical narrow, hummocky rim deposits, radial ridges, and surrounding extensive field of secondary craters. Interior terraces and central peaks are well-developed and typical of mercurian craters of this size. Smaller craters in foreground, about 25 km diameter, are also terraced. After Gault et al. (1975). (c) Two craters, 128 and 195 km in diameter, that have interior rings of mountains and ejecta deposits deeply scarred by chains of secondary craters. (d) Changes in interior structure of mercurian craters and inner ring diameter as a function of the rim diameter. From Gault et al. (1975).
concentric with—and approximately one half the diameter of—the main rim of the crater (Figure 5c and 5d). This progressive change in crater geometry also occurs on the Moon, but the changes from one morphologic type to another occur at smaller diameters on Mercury. Transition at smaller diameters on Mercury is consistent with differences in gravitational acceleration if the changes in morphologic type are caused by gravity-induced modifications (collapse) of the original craters of excavation (Gault et al. 1975).

The youngest craters (sharpest morphology), especially those near the limb under conditions approaching zero-phase lighting, display well-developed, extensive bright-rayed ejecta systems extending hundreds of kilometers and a few cases extending more than a thousand kilometers from their parent crater (Figure 1b). Some of these rayed craters are also centered within dark halos comprised of low albedo material. Although such prominent rays of ejecta and dark halos are common to lunar craters, it is the morphology of the main ejecta deposits on the flanks of the crater rims that distinguishes mercurian craters from their lunar cousins. For both families of craters the rim deposits consist of a hummocky facies that grades out with increasing distance from the rim into a radially ridged facies which, in turn, changes into a zone of satellitic secondary craters and discontinuous ejecta deposits (Figure 5b). On Mercury, the hummocky and radially ridged facies—the ejecta deposits termed the continuous deposits—grade into the satellitic crater field over a much shorter distance than on the Moon. Additionally, the areal density of the craters in the satellitic zone is very high compared to lunar secondary crater fields. Chains of overlapping craters resulting in long, linear grooves (subradial to the primary crater) extend across the continuous deposits almost up to the crater rim (Figure 5c), a characteristic that is rare on the Moon. The narrower continuous deposits, linear grooves, and increased areal density of the mercurian secondary craters relative to lunar craters is totally consistent with the reduced ballistic range of ejecta caused by the differences in the gravity environments between the two planetary bodies.

Size-frequency distributions of mercurian craters have been obtained as a basis for determining relative ages of the major physiographic provinces and correlating interplanetary geologic time (Murray et al. 1974, Guest & Gault 1976). The cratered terrain, comprised of both the intercrater plains and heavily cratered terrain provinces, has a crater size-frequency distribution effectively the same as for the southern highlands on the front side of the Moon. The two provinces were combined for purposes of crater counting because it is difficult to separate the two provinces into discrete areas (Guest & Gault 1976), although it is recognized that they may represent two different epochs in Mercury's history. Frequency distributions for the mercurian smooth plains are greatly reduced relative to the cratered terrains and are almost identical to those for the Apollo 14 landing site in the Fra Mauro formation. Crystallization ages of rocks from this Apollo site have been dated to be about $3.9 \times 10^9$ years (Papanastassiou & Wasserburg 1971, Turner et al. 1971), but the similarity in the mercurian smooth plains and Apollo 14 crater populations does not equate directly to similar absolute ages. Differences between the Moon and Mercury in their gravity environments, average impact
velocities, and the sources (i.e. time-integral fluxes) of the bodies forming the craters have a direct effect on the resultant frequency distributions that precludes dating in absolute terms. The sources or origins of the crater-forming bodies is a subject of particular significance in the geologic history and are discussed briefly later; it is sufficient here to note that the mercurian heavily cratered terrains constitute the oldest surfaces on Mercury and if analogous to the lunar highlands must be older than 4 billion years (Murray et al. 1975).

Guest & Gault (1976) have examined the morphology of the mercurian crater populations and find that all young, morphologically sharp mercurian structures smaller than about 30 km must have been formed subsequent to the formation of the Caloris basin. Their results suggest that either some process or event degraded such craters prior to (or contemporaneously with) the formation of Caloris or, possibly, such smaller, sharp craters were never formed in the time period before the Caloris event; the observations cannot distinguish between the two alternatives and the question remains open to further investigations.

Optical and Thermal Properties

Mercury's optical and thermal emission properties (normal albedo, photometric function, polarization, spectral characteristics, thermal inertia, etc.) have been measured from Earth and are virtually identical to those of the Moon (Lyot 1929, Harris 1961, de Vaucouleurs 1964, Irvine et al. 1968, Morrison 1970, McCord & Adams 1972a,b, Dollfus & Auriere 1974, Vilas & McCord 1976, Chase et al. 1974, 1976). Observations by Mariner 10 have supplemented these Earth-based data of the integral disk in addition to providing results in the visible for surface resolutions of 20 km (Murray et al. 1974, Hapke et al. 1975) and 40-90 km in the far infrared (Chase et al. 1974, 1976).

Normal albedos at 0.554-μm wavelength were obtained by Mariner 10 for large areas of the incoming and outgoing hemispheres (Figure 1). Values of 0.16 and 0.12 were found (Hapke et al. 1975) for, respectively, areas consisting primarily of intercrater plains and smooth plains with an average value of about 0.14. For comparison, the recent polarimetric determinations by Dollfus & Auriere (1974) for the integral disk give 0.13. Although phase-angle observations by Mariner 10 were restricted to a narrow range because of trajectory and camera pointing limitations, relative brightness curves across the illuminated disk are in excellent agreement with typical lunar curves for the same phase angles. Photometrically, therefore, as major physiographic provinces, the brighter intercrater plains resemble the lunar highlands and the darker smooth plains resemble the lunar maria. Albedos vary within these provinces from 0.09 to 0.21 with the suggestion that there are two types of smooth plains: one with an albedo of about 0.13 within and around the Caloris basin and a second with an albedo of around 0.20 found in both the incoming and outgoing hemispheres of Figure 1. Mercurian rayed craters are generally brighter than lunar rayed craters with some albedos as high as 0.45 (Hapke et al. 1975). With the exception of these bright features there are no large albedo gradients on the planet comparable to the lunar highland-maria contrasts.

Polarization measurements of Mercury, first performed by Lyot (1929), are
recently summarized by Dollfus & Auriere (1974). Measurements obtained for different areas of the planet and for the integral disk at different wavelengths are virtually identical to corresponding Earth-based measurements for the Moon. In particular, the negative branch of the polarization curves is indicative of the presence of a dark, fine-grained powder similar to the lunar regolith on all of Mercury's surface. Mariner 10 polarimetric observations (Hapke et al. 1975) with a surface resolution of 20 km failed to detect any anomalously high polarization that would be indicative of landforms not covered by a regolith-type soil. Consistent with the absence of large albedo gradients, variations in polarization across the surface are minimal, less than 0.02 within the value of 0.12.

Evidence for mineralogy and composition of surface materials is contained in spectral reflectance measurements in the wavelength region from about 0.3 to 2.0 μm (e.g. Adams 1968, 1974, 1975). McCord & Adams (1972a, b), using a 22-channel system operating from 0.32 to 1.05 μm, found a constant slope for the reflection spectrum of the integral disk with the exception of a weak absorption band near 0.95 μm common to pyroxenes. Because the spectral curve matches closely that for the lunar highlands and maria, they conclude that Mercury's surface is mantled with a lunar-like regolith of iron and titanium rich glasses. More recent observations with refined instrumentation (Vilas & McCord 1976) essentially re-affirm the earlier results; some differences are noted in the spectra that could be attributed to differences in composition for the two different regions of Mercury observed.

Although the Mariner 10 polarization pictures are bland, the planet's surface reveals considerable color variations (Hapke et al. 1975). The color images were formed as composites of the same scenes taken through filters having effective wavelengths of 0.355 and 0.725 μm. Areas with the same reflectance spectrum appear uniformly grey; areas of redder or bluer materials appear, respectively, as brighter or darker features. These "color" pictures reveal that all fresh bright-rayed craters are consistently bluer than their surroundings. No red-rayed craters were found, although most other bright areas generally were redder in the same manner as on the Moon. The unusual combination of high albedo and bluish color of the rayed craters suggests to Hapke et al. (1975) that, in marked contrast to the conclusions of McCord & Adams (1972a, b), the mercurian crustal material is low in titanium and iron (Fe+2). The difference between the two analyses remains unresolved, but the difference may be real because the two data sets do not correspond to the same area of the planet; the Earth-based data correspond to surfaces that were behind the terminators during the Mariner 10 flybys.

Observations at longer wavelengths in the far infrared and microwave regions reinforce the lunar-like characteristics of the mercurian surface materials (e.g. Morrison 1970). Most recently, Chase et al. (1974, 1976) employed a two-channel infrared radiometer (11 and 45 μm) on Mariner 10 to scan the dark side of the planet along an approximately equatorial band with a 40- to 90-km resolution. Except for a region of strong local enhancement, values of the thermal inertia \( k_{pc}^{-1/2} \) ranged from about 0.0016 to 0.0026 cal cm\(^{-2}\) sec\(^{-1/2}\) K\(^{-1}\) (\( k \) is thermal conductivity, \( pc \) is heat capacity/volume). Such values together with derived thermal
skin depths for diurnal variations, electrical skin depth, and loss tangent all fall within the range of values for the Moon.

It is interesting that the highest values of the thermal inertia occurred in the quadrant behind the morning terminator. This seems compatible with a young surface that presumably includes the hidden half of the Caloris basin structure (and external smooth plains) if it has circular symmetry. The largest values of thermal inertia (0.0031) occur in a region of local enhancement 400 km wide, part of which is within an area probed, but not evident, by radar (Zohar & Goldstein 1974). However, a second region of enhancement (40 km) appears to correlate with a similar-sized region of strong radar backscatter, suggesting a young crater surrounded by a rocky ejecta deposit. Two additional anomalies lay outside radar coverage, one of which could be another young crater.

Microwave radiometry samples of the integral mercurian surface to meter depths are in substantial agreement with observations at the shorter wavelengths; data for Mercury closely approximate that which would be expected for the Moon if it were placed in a mercurian orbit. The reader is referred to an excellent review and analysis of the literature by Morrison (1970) (see also Devine 1972). More recent microwave measurements and refined analyses are reported by Morrison & Klein (1970), Ulich, Cogdell & Davis (1973), Briggs & Drake (1973), and Cuzzi (1974).

Cartography

Acquisition of the Mariner 10 imagery has initiated a program of high-resolution mapping of Mercury (Davies & Batson 1975). The coordinate system used in this effort assumes that the equator lies in the plane of the orbit and the center of a small crater (named Hun Kal, see Figure 5a and below) defines the 20° meridian; longitude is measured from 0° to 360°, increasing to the west. For comparison, the IAU (see Davies & Batson 1975) define 0° longitude to be the subsolar meridian at the first perihelion after January 1, 1950 (Julian date 2,433,292.63) with the spin axis normal to the orbital planet. The Mariner 10 system places the IAU zero meridian at approximately 359.5° with an uncertainty of about 0.5°. The use of a body-fixed (crater) criterion for longitude is advantageous in preserving the validity of the relative positions of features on the maps until future exploration permits relating the two systems with greater accuracy than is currently possible (20 km).

The cartographic products consist of 1:5,000,000 shaded relief maps prepared by the U.S. Geological Survey, Flagstaff, Arizona. The surface of Mercury in this series is divided into 15 areas: five Mercator projections encircling the planet between 25° north and south latitudes; eight Lambert projections between 20° and 70° latitudes; and two polar stereographic projections from 65° to the poles.

Nomenclature for the topographic features on the maps, as established by the IAU (Morrison 1976), comprises six major topographic features: craters, valleys, scarps, ridges, plains, and mountains. Craters are named for authors, artists, and musicians (e.g. Homer, Renoir, Bach). Two exceptions are Kuiper and Hun Kal; the former commemorates Dr. Gerard P. Kuiper, who was a member of the Mariner 10 Imaging Team and who died prior to the spacecraft's encounters with Venus and Mercury; the latter (the number twenty in ancient Mayan) defines the
MERCURY

20° meridian. Valleys (vallis) are named after radio observatories, and scarps (rupe) are given names of ships associated with exploration and scientific research. Ridges (dorsum) are not named after any specific group, but plains (planitia) are assigned names for the word meaning Mercury (planet or god) in various languages.

INTERIOR

Composition and Structure

The mass and radius determinations imply that Mercury must contain a large fraction of iron—the only heavy element sufficiently abundant to account for the high density. As inferred from solar abundances, and by analogy with terrestrial, meteoritic, and lunar materials, it is presumed that silicates of iron and magnesium are also prevalent. The similarity of surface morphology to lunar counterparts suggests (see below) a mercurian mantle of differentiated silicates at least grossly like the Moon in composition. This view is borne out by ground-based and Mariner 10 observations which indicate that the radio and optical (including IR) properties of Mercury's surface are essentially the same as those of the lunar surface (Harris 1961, de Vaucouleurs 1964, Hameen-Antilla, Pikkarainen & Camichel 1970, Dollfus & Auriere 1974 and references therein, McCord & Adams 1972a, b, Murray et al. 1974, Hapke et al. 1975, Vilas & McCord 1976).

Bounds on the amount of iron present can be determined by solving the spherically symmetric hydrostatic equation,

\[ \frac{dP}{dr} = - \frac{\rho G m(r)}{r^2} \]

with an appropriate equation of state, \( P = P(\rho, T) \), and requiring the solutions to conform to the observed mass and radius of the planet. (\( P \) is the pressure, \( \rho \) the density, \( T \) the temperature, \( G \) the gravitational constant, and \( m \) the mass within radius \( r \).) The solutions, of course, are not unique. The equation of state depends on the temperature, but for solids and liquids at planetary pressures, thermal expansion is generally less than the volume change due to compressions, so that even constant temperature models are good approximations. More important are the extent to which the planet is differentiated (specifically, how much iron has separated into a core) and possible phase changes, both of which affect the equation of state to be used at a given radius. Detailed models of Mercury have been calculated by Reynolds & Summers (1969) which include two extreme hypotheses: 1. all iron is concentrated in a core; and 2. all iron is in oxide form distributed homogeneously throughout the planet. They showed that the overall abundance of iron is rather insensitive to assumptions regarding its distribution, a result confirmed by the calculations of Siegfried & Solomon (1974) (see also Kozlavsksaya 1969). The mass fractions of total iron for all models fall within the range 0.60-0.71, as compared to about 0.34-0.38 for the Earth. The density and pressure are shown in Figure 6 for two core models. The radius of a fully differentiated core is expected to be approximately three fourths the planetary radius; this result is insensitive to assumptions regarding the detailed composition of the mantle. The central pressure does not exceed 500 kbar. In a homogeneous model it may be as low as 245 kbar.

Lewis (1972) has proposed a model of the chemical compositions of the planets based on the assumptions that: 1. equilibrium chemistry prevailed as the planetary
material condensed from a hot solar nebula; 2. the structure of the nebula was similar to that proposed by Cameron (1969), particularly with regard to the variation of temperature with heliocentric distance; and 3. the material which did not condense was somehow removed from the vicinity of the planet before the nebula completely cooled. This model is popular because it provides a simple explanation for the mean planetary densities. For Mercury, Lewis predicts (besides an enrichment of iron relative to magnesium and silicon with respect to solar abundances) strong depletion of sulfur, alkali metals, and volatiles. At present, our knowledge of Mercury's composition is insufficient to confirm or refute such a model, except with regard to the enhanced iron abundance.

Thermal History

The principal constraints on Mercury's thermal history drawn from the Mariner 10 observations (see section on geologic history below) are the following: 1. the surface of the planet is the product of extensive differentiation, which implies that at least the outer portion of the planet melted; 2. the magnetic field is intrinsic to the planet, and is most likely associated with an internal dynamo (either extant or extinct), which implies global differentiation involving the formation of an iron core; 3. differentiation occurred early in the planet's history (probably within the first 1/2 billion

![Figure 6](image)

Figure 6  Pressure and density as functions of radius for two models of a differentiated Mercury. Differences are due primarily to different assumptions regarding the equations of state.
years); and 4. the surface of the planet has been largely undisturbed by tectonic
processes for most of the time since differentiation. In addition, Solomon (1976) has
argued that if a molten core existed at one time, it could not have completely
solidified since the cessation of surface activity without resulting in contraction of the
planet much greater than that which has been inferred from the compressional
features described by Strom, Trask & Guest (1975).

What could be the source of energy for melting much of the planet earlier than
4 billion years ago? Mercury may have formed from material already at a relatively
high temperature due to its proximity to the Sun; Lewis (1972) proposed that the
planet is composed only of substances that would be condensed at 1100°C. In
addition to this internal energy, gravitational energy equivalent to an average tem-
perature increase of 10^3 °C, if all of it was retained, is released upon formation of
the planet. However, much of this energy would be radiated to space. Models of the
accretion process (Benfield 1950, Mizutani et al. 1972, Weidenschilling 1974) show
that the amount of energy retained by the planet depends on the rate of energy
deposition during formation: rapidly forming bodies retain more heat. Also, the
accretional energy is deposited nonuniformly, with most of it going into the outer
half of the planet. Siegfried & Solomon (1974) report that models based on the
ad hoc mass accretion law of Hanks & Anderson (1969), and with an initial tem-
perature of 1127°C, radiate away almost all of the gravitational energy if the
accretion time is longer than 10^5 years. More detailed (but not necessarily more
realistic) models based on the theory of the gravitational capture of small objects
would retain more heat for the same accretion time. But recent work by Weiden-
schilling (1974, 1976) favors Safronov’s (1969) estimate of accretion times for the
terrestrial planets of ~ 10^8 years. Such slow accretion would yield negligible heating.

Accretional energy is deposited at the surface of a growing planet, and is therefore
lost easily: energy deposited deep within a terrestrial planet is not lost quickly
because rocks are good insulators. (In the absence of convection, the characteristic
thermal transport time is greater than the age of the solar system.) It is well known
that the Earth and Moon are heated from within by small amounts of long-lived
radioactive elements – primarily 232U, 235U, 232Th, and 40K, which have half-lives
on the order of 10^9-10^10 years. The minimum heat source density H required to
melt the center of Mercury in 1/2 billion years can be estimated by assuming that
no heat escapes the region during that time. For a temperature increase of 2,000°C,
H = 8 × 10^-6 ergs cm^-3. This rate of heat generation is three times that which
would have existed in chondritic meteorites 4.6 × 10^9 years ago, and four times
that for the uranium-enriched, potassium-depleted Mercury implied by Lewis’ model.
Thus, if these radioactive elements provided the energy for early differentiation,
their abundances must have substantially exceeded those found in other solar system
objects, as well as current theoretical estimates.

Relatively short-lived isotopes – notably 26Al (half-life = 0.7 × 10^6 years) have
also been suggested as important heat sources (Urey 1955a). Evidence for the
existence of 26Al in the early solar system has recently been discovered by Lee et al.
(1976), who find a magnesium isotopic anomaly in a sample of the Allende meteorite
which is most plausibly explained by the in situ decay of 26Al. They point out
that even a body of asteroidal size would be heated to melting if it contained
the abundance of $^{26}\text{Al}$ inferred for their sample. Although the abundance
and distribution of this isotope in early solar system bodies remains a matter for further
study, it must be considered as a possible energy source for the differentiation of
Mercury. (Schramm 1971 discusses the theoretical problems associated with $^{26}\text{Al}$ in
the early solar system.)

Other sources of heating are the dissipation of tides and the extraction of energy
from an early solar wind by electromagnetic induction (Sonett et al. 1968). The
former has been evaluated by Burns (1976) and found to be insufficient to produce
global differentiation. The latter mechanism (besides requiring the Sun to go through
a T-Tauri-like phase of rapid rotation with a high magnetic field) requires that
a good electrical contact be maintained between the planet and the solar wind
during the heating. It is not known if these conditions existed. Under the most
favorable conditions, the induction mechanism appears to be capable of raising the
internal temperature of Mercury by more than $10^3\,^\circ\text{C}$. (Sonett et al. 1968).

If a molten core did form early in Mercury's history, how long could it remain molten? Thermal history models of Siegfried & Solomon (1974) and Solomon (1976)
indicate that an iron (or iron-nickel) core would solidify in less than 2 billion years.
But the rate at which heat leaves the core is sensitive to conditions at the core-
mantle boundary. In the above mentioned models (and those of Sharpe & Strangway
1976), the melting temperature is taken to be continuous and superadiabatic at
the boundary. This assumption has the consequence that, as long as the core is
molten at the boundary, the heat flux there must be at least $k(dT/dr)_{\text{melting}}$ which,
when integrated over the boundary surface, removes energy from the core at the
rate of $1.6 \times 10^{19}\,\text{ergs sec}^{-1}$. (The thermal conductivity of iron is denoted by $k$;
see Liu & Bassett 1975 for the melting curve of iron at the boundary pressure
70 kbar.) On the basis of geochemical arguments (Urey 1955b), which are borne
out by the low radioactive abundances found in iron meteorites (Mason 1971), it is
assumed that radioactives do not exist in significant amounts in the core, and
therefore cannot contribute to this power. If it is to be supplied solely by heat of
fusion, solidification must occur in $1.7 \times 10^9\,\text{years}$, as indeed occurs in the models.

However, Fricker et al. (1976) have pointed out that the melting temperatures
of silicates are likely to be greater than that of iron at the core-mantle boundary.
They have calculated an evolutionary model in which there is a discontinuous drop
in melting temperature (of about 200°C) from the silicate mantle to the iron core.
The discontinuity has an important effect: the temperature in the outer core
increases above the melting point, while the gradient decreases, thereby reducing the
heat flow from the core. In their model, a molten shell occupying the outer 500 km
of core is retained to the present. The thermal gradient in the shell is subadiabatic,
so the model could not explain a convectively driven dynamo. If it is required
that the core be convective, the heat flux from the convecting region must be at
least $k(dT/dr)_{\text{adiabatic}}$. Using recent data on the properties of iron at high pressure
(G. C. Kennedy, personal communication, 1977, Liu & Bassett 1975), and models
of Mercury's structure (Reynolds & Summers 1969), we find that the adiabatic
gradient at the core-mantle boundary is about $1^\circ\text{C} \,\text{km}^{-1}$—near the melting curve
gradient!—so that, again, the heat of fusion would be used up in less than 2 billion years if it were the only available source.

Another consequence of a convecting core is that the lower mantle must either be molten or solid convecting. In the absence of mantle convection, the greatest heat flow that can be expected to occur in the silicate material at the core-mantle boundary is that which would correspond to the conductive, steady-state temperature distribution with no heat sources in the mantle. If the silicate melting temperature is not too different from that of iron, this heat flow is only $6.5 \times 10^{18}$ ergs sec$^{-1}$ (Cassen et al. 1976) or less than half the minimum convective flux from the core. (This disparity is a consequence of the fact that the thermal conductivity of the silicates is typically an order of magnitude less than that of iron.) The mantle must either heat up to melting, where liquid convection can carry off the heat, or convect in the solid state. Cassen et al. (1976) show that the latter is likely. In fact, mantle convection is probable even if the core is not convecting, as long as the core-mantle boundary is near the iron melting point, and the solid viscosity is similar to that of the Earth's upper mantle. This conclusion, however, provides a further disadvantage for the convecting dynamo hypothesis; mantle convection could extract as much as $3 \times 10^{19}$ ergs sec$^{-1}$ from the core, causing it to solidify in less than a billion years. Retention of heat sources throughout the mantle in an amount exceeding that found in chondritic meteorites (or that inferred for the Earth) would be necessary for the maintenance of a molten (but nonconvection) core.

In view of the difficulty of explaining the source of energy lost by a convective dynamo, it seems worthwhile to pursue Stephenson's (1976) residual dipole model. Although an iron core might solidify rapidly, it is very unlikely that it could cool below the Curie point; therefore, the core would be paramagnetic. In this case, Stephenson finds that about $3\%$, (by volume) or more metallic iron would probably be needed in that part of the mantle which cooled below the Curie point while the dynamo was operating. This assumes that remanence is induced in a shell somewhat greater than 200 km thick, the most that could be hoped for according to thermal history calculations. Thinner regions would require more free iron; any overturning, local reheating, or other disruption of the region would diminish the resultant dipole moment. Lunar basalts, which carry their remanence in metallic iron grains, generally contain two orders of magnitude less free iron than that suggested for Mercury. It is an open question whether or not Mercury—an iron-rich planet—might have more iron in its mantle than is found in the lunar rocks.

GEOLOGIC HISTORY

Despite some important differences, the striking duplication between the Moon and Mercury in their surface features and relative relationships suggests that the sequence of geologic events leading to their present state must have been very similar. Whether the absolute time scale for development of the surface features of Mercury (and other terrestrial planets) is the same as for the Moon remains an open question. Within the limitations of the unknown time scale, Mercury's evolution can be divided into five stages or epochs: 1. accretion and differentiation; 2. terminal
heavy bombardment; 3. Caloris basin formation; 4. basin flooding; and 5. post-
filling light bombardment (Murray et al. 1975).

The first epoch includes the earliest stages of the solar system, beginning with
the condensation of the solar nebula into solids and the accumulation of the solids
into the main mass of Mercury. The details of this accumulation are not known,
but the evidence for large-scale volcanism of silicate composition grossly similar
to the Moon, and the existence of an iron core inferred from the magnetic field
observations, argue persuasively that Mercury is a chemically differentiated planet.
This segregation into an iron core-silicate crust must have occurred very early
during this epoch in order to form a lithosphere of sufficient rigidity and thickness
to permit the heavily cratered terrain (formed in the next epoch) to retain its lunar
appearance. Because of the absence of any evidence for colian erosion, Murray et al.
(1975) argue that any atmosphere formed during this period must have been
dissipated very quickly or before the onset of the terminal heavy bombardment.
Chapman (1976) disputes this interpretation from considerations of the relative rates
of erosion and cratering; by inference he implies any primordial mercurian atmos-
phere may have been retained into the second epoch of heavy cratering.

The heavily cratered terrain, represented by large craters several tens of kilometers
in diameter grading up to basins the size of Caloris, records a period of intense
or prolonged bombardment by large bodies. This terrain appears to be emplaced
into an older host material, the intercrater plains whose origin, as previously dis-
cussed, is very uncertain. Because of the uncertainty, a point of discussion has arisen
whether the bombardment producing the heavily cratered terrain was the terminal
phase of the accumulation of Mercury, or whether it was a second episode of
 cratering not related to accretion. Murray et al. (1975) interpret the extensive areas
of intercrater plains to be degraded remnants of all topographic evidence of the
saturation flux that necessarily constituted final stages of accretion. This interpreta-
tion has an important implication: if the intercrater plains are the degraded
accretionary stage, then the heavily cratered terrain must record a later episode of
bombardment distinct and separate from accretion.

A similar scenario has been proposed for the Moon to explain a rather sharp
cutoff in crystallization ages of lunar samples at about 4 billion years (Tera,
Papanastassiou & Wasserburg 1974). The proposal involves a “cataclysmic event”
about 3.9 billion years ago, an event of about 10^8 years' duration (including the
Imbrium and Orientale basin-forming events), and is conceived as a spike on a curve
of the cratering flux that was decreasing from the high rates that occurred during
accretion. Baldwin (1974), Hartmann (1975), and Chapman (1976) have pointed out
the poor definition of the proposed cataclysm, but the latter also points out that
the evidence does not exclude a possible cataclysm. Indeed, recent studies of lunar
 crater morphologies (Whitaker & Strom 1976) have been interpreted to indicate
that two different populations of bodies have cratered the Moon on the appropriate
time scales. Two sources for the “cataclysmic” flux have been given, both of which
could have arisen as the natural consequences of the dynamical evolution of the
early solar system. Wetherill (1976) has proposed that a large planetesimal formed
in the vicinity of Uranus and Neptune was perturbed on a time scale of several
10^8 years into an Earth-crossing orbit, and during a close approach to either Earth or Venus was tidally disrupted. The subsequent sweepup of the fragments by the inner planets produced the “cataclysm.” Chapman & Davis (1975) have suggested an alternate source from the asteroid belt; the parent body, in a highly eccentric orbit, was disrupted by collision to provide the population of fragments for the terminal heavy bombardment. In either case, as Chapman (1976) points out, if an episodic-type of cataclysmic bombardment did occur on the Moon, it must also have occurred on Mercury and the other terrestrial planets at the same time of about 4 billion years ago. Preliminary studies of mercurian crater morphology (Guest & Gault 1976) and total crater populations (Whitaker & Strom 1976) are suggestive of two families of impacting bodies in support of the Murray et al. (1975) interpretation of an episodic late heavy bombardment. However, other interpretations of the available mercurian crater data are equally plausible, and the heavily cratered terrain does not necessarily need to be interpreted as part of an inner solar system “cataclysmic” bombardment; bodies in initially almost circular Mercury-crossing orbits or in eccentric orbits with perihelia well inside Mercury’s orbit would have had cratering rates on Mercury an order of magnitude greater than on the other terrestrial planets (Wetherill 1974). Further studies of Mariner 10 imagery, especially the mercurian craters, may be able to resolve the uncertainties of an episodic bombardment versus the tail end of accretion as the source of the late heavy bombardment and the heavily cratered terrain.

An important aspect of the intercrater plains and heavily cratered terrain that differs from the lunar surface is the presence of the lobate scarps of probable tectonic origin. These features appear predominantly on the intercrater plains, a pattern of occurrence that suggests that formation of the scarps was primarily during the latter stages of the first epoch, and perhaps continued into the early stages of the late heavy bombardment that produced the heavily cratered terrain. If these scarps are thrust or reverse faults arising from compressive stresses and consequent crustal shortening caused by cooling and shrinkage of the core (Strom, Trask & Guest 1975), the scarps represent an important boundary condition for thermal history calculations once an absolute time scale for intercrater plains and heavily cratered terrain can be established.

A convenient and well-delineated third point in Mercury’s history is the time of the impact that formed the Caloris basin. This monstrous event modified a major fraction of the “outgoing” hemisphere in a manner similar to that forming the ringed basins on the Moon (Strom, Trask & Guest 1975). In addition to direct modifications by excavation and deposition processes, volcanic activity may have been triggered that is responsible for some of the elements of the smooth plains. Stratigraphic relationships, areal density of crater populations, and crater morphologies indicate that the short epoch of the formation of Caloris either occurred very near to, or marks the end of, the late heavy bombardment (Trask & Guest 1975, Guest & Gault 1976).

The fourth period of mercurian history started an indeterminate but probably very short interval of time after the Caloris impact. Broad expanses of smooth plains were formed and smaller, older basins were filled with similar appearing
materials, most plausibly by volcanic origins (Trask & Guest 1975, Strom, Trask & Guest 1975, Trask & Strom 1976). The low areal density of craters on the smooth plains clearly indicates that the plains post date the late heavy bombardment. Small variations in the areal densities between different units of the smooth plains emphasize that the filling was not simultaneous on a global scale; although the time required for plains formation is presently uncertain, the interval was probably relatively short (Murray et al. 1975).

The fifth and final epoch that can be defined for Mercury extends from the termination of basin flooding to the present, probably a major fraction of mercurian history. It was a quiescent period during which the surface acquired a light “dusting” of meteoritic debris which produced most, if not all, of the craters now expressed on the smooth plains and the prominent rayed craters, the youngest features displayed on the surface. Degradation of craters and presumably other landforms has been minimal during this period (Guest & Gault 1976), and no tectonic deformations of the mercurian crust are evident (Murray et al. 1975). Like the Moon, Mercury currently appears to be resting quietly with no active processes to reshape its surface except an occasional random encounter with yet another fragment of interplanetary debris.

Acknowledgments

The authors express their appreciation to A. L. Broadfoot, D. E. Shemansky, R. T. Reynolds, S. J. Peale, R. Greeley, and P. Dyal for helpful reviews and discussion. A special thanks to Faye L. Gray for preparing the typescript to meet an impossible deadline.

Literature Cited

Ogilvie, K. W., Scudder, J. D., Hartle, R. E., Siscoe, G. L., Bridge, H. S., Lazarus, A. J.


