The Earth-Moon Connection

A. E. Ringwood
Research School of Earth Sciences Australian National University Canberra, Australia

Z. Naturforsch. 44a, 891-923 (1989); received June 15, 1989

Dedicated to Heinrich Wänke on the occasion of his 60th birthday

The early thermal state of the Earth provides important constraints on hypotheses relating to its origin and its connection with the Moon. The currently popular giant impact hypothesis of lunar origin requires the Earth’s mantle to have been completely melted during the impact. Differentiation of a molten mantle would have produced strong chemical and mineralogical stratification, causing the mantle to become gravitationally stable and resistant to convective rehomogenization. The resulting composition and mineralogy of the upper mantle and primitive crust would have been dramatically different from those which have existed during the past 3.8 by. It is concluded that the Earth’s mantle was not extensively melted at the conclusion of accretion of the planet and therefore the hypothesis that the Moon was formed by the impact of a martian-sized planetesimal on the proto-Earth is probably incorrect. Nevertheless, a wide range of geochemical evidence demonstrates the existence of a close genetic relationship between the Moon and the Earth’s mantle. The key evidence relates to the processes of core formation in planetary bodies and resultant abundance patterns of siderophile elements which remain in their silicate mantles. Because of the complexity of the core formation process within a given body and the multiplicity of chemical and physical processes involved, the mantle siderophile signature is expected to be a unique characteristic. Thus, the siderophile signatures of Mars and of the eucrite parent body are quite distinct from that of the Earth’s mantle. Lunar siderophile geochemistry is reviewed in detail. It is demonstrated that a large group of siderophile elements display similar abundances in the terrestrial and lunar mantles. The similarity implies that a major proportion of the material now in the Moon was derived from the Earth’s mantle after core formation. This implication, however, does not require that the bulk compositions of the lunar and terrestrial mantles should be essentially identical, as is often assumed. Factors which may contribute to significant compositional differences between the two bodies within the context of a close genetic relationship are reviewed. The most promising mechanism for removing terrestrial material from the Earth’s mantle arises from the impacts of a number of large (0.001 to 0.01 Mn) but not giant (≥ 0.1 Mn) planetesimals after core formation and at the terminal stage of the Earth’s accretion. These impacts evaporated several times their own masses of mantle material and considerably more. However, they did not lead to complete or extensive (e.g., > 50%) melting of the entire mantle. Impact-generated clouds of shock-melted spray and vapours were accelerated to high velocities in the presence of a primitive terrestrial atmosphere that co-rotated with the Earth. This provided an effective means of transferring angular momentum from the Earth to the ejected material which condensed to form a ring of Earth-orbiting planetesimals and moonlets. The Moon was formed by coagulation from material derived from the outer regions of this ring. Accretion of the Earth in the presence of the gases of the solar nebula and the co-rotating primitive terrestrial atmosphere may also have provided a mechanism for generating the rapid prograde spin of the proto-Earth.

1. Introduction

I am delighted to help celebrate the 60th birthday of Heinrich Wänke, a warm friend and respected colleague for many years. It seems appropriate under these circumstances to contribute a paper on a subject which is dear to his heart and to which he has made so many major contributions. Arguably, the most fundamental problem in the earth sciences relates to the origin of this planet. It is now widely recognized that this problem, in turn, is inextricably linked to the origin of the Moon. Heinrich pioneered in the discovery of a compelling array of geochemical evidence which points to the close genetic relationship between these two bodies. The existence of this genetic relationship is now widely accepted. Somewhat ironically, acceptance has been achieved more by new dynamical arguments rather than the earlier geochemical evidence. The dynamical arguments have led to the "giant impact" hypothesis of lunar origin which recently seems to have achieved bandwagon status. Without doubt, the...
new dynamical concepts are of great importance and represent a major advance. Nevertheless, I believe that the currently popular version of the giant impact hypothesis is faced by certain problems which have not been adequately recognized and that it is the earlier geochemical history which actually provides the most decisive evidence for the genetic relationship between Earth and Moon.

The early thermal state of the Earth provides important constraints on the issues outlined above. Whether the Earth began as a hot, molten body, as a cool, unmelted and undifferentiated planet, or in some intermediate state, was strongly influenced by the manner of its accretion within the primordial solar nebula. Accretion is believed to have been a hierarchical process in which dust particles coagulated to form small planetesimals which in turn accreted to form larger, intermediate-sized bodies that were ultimately assembled to form the Earth. The process involved a complex equilibrium between planetesimal growth, collisions and fragmentation.

Modern cosmogonic models display widely divergent views concerning the size distribution of planetesimals and intermediate-sized bodies. According to an elegant scenario developed by Wetherill [1], planetesimals grew to form very large intermediate-sized bodies before finally combining to form the Earth. His model proposes that 1–2 martian-sized bodies before finally combining to form the Earth. His model proposes that 1–2 martian-sized bodies before finally combining to form the Earth. The process involved a complex equilibrium between planetesimal growth, collisions and fragmentation.

Modern cosmogonic models display widely divergent views concerning the size distribution of planetesimals and intermediate-sized bodies. According to an elegant scenario developed by Wetherill [1], planetesimals grew to form very large intermediate-sized bodies before finally combining to form the Earth. The process involved a complex equilibrium between planetesimal growth, collisions and fragmentation.

The above considerations suggest that if we could establish some constraints upon the Earth’s initial thermal regime and degree of differentiation, these in turn might provide insights into the nature of the accretion process and could also help to evaluate the giant impact model for the origin of the Moon.

2. Constraints on the Melting and Early Differentiation of the Earth

(a) Early Thermal State

During the 1950’s and 1960’s, it was almost universally believed that the Earth had accreted in a relatively cool, unmelted state, and that it was then heated by radioactive decay of U, Th and K over a period of a billion years or so until temperatures were high enough to permit core segregation, which occurred over an extended period [14–23]. However, Ringwood [24–26] pointed out that core-segregation would have substantially altered the U/Pb ratio of the mantle-crust system. The “age of the Earth” of 4.5 b.y. which had been obtained by Patterson [27] from the lead isotopic systematics of crustal rocks must therefore date the time of core formation. Ringwood concluded that the Earth had accreted in a relatively hot condition, to enable core formation to occur during or very soon after accretion. This conclusion was further supported by measurements of the partition coefficients of lead between iron alloys and silicates which showed that lead would have been preferentially partitioned into the core, as compared to uranium [28].

A study [29] of the early thermal evolution of the Earth suggested that a magma ocean about 200 km deep was likely to have been formed immediately following the completion of accretion. Partial melting
Table 1. Partition coefficients for selected elements between liquidus MgSiO₃ perovskite, pyropic garnets and ultrabasic melts. From Kato et al. [36]*.

<table>
<thead>
<tr>
<th>Element</th>
<th>MgSiO₃(pv) D_{mpv/liq}</th>
<th>Garnet D_{gnt/liq}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ca</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Al</td>
<td>0.5–0.8</td>
<td>2.5</td>
</tr>
<tr>
<td>Na</td>
<td>0.02</td>
<td>0.1</td>
</tr>
<tr>
<td>Sc</td>
<td>5</td>
<td>1.7</td>
</tr>
<tr>
<td>Y</td>
<td>3</td>
<td>1.3</td>
</tr>
<tr>
<td>Yb</td>
<td>2</td>
<td>1.4</td>
</tr>
<tr>
<td>Sm</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>La</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>Ti</td>
<td>3</td>
<td>0.4</td>
</tr>
<tr>
<td>Hf</td>
<td>14</td>
<td>0.8</td>
</tr>
<tr>
<td>Zr</td>
<td>9</td>
<td>0.6</td>
</tr>
<tr>
<td>Nb</td>
<td>~1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>K, Sr, Ba, Rb</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>

* This is the preferred set of partition coefficients obtained in the investigations of Kato et al. [36]. It is believed that they are likely to be correct to ±50%, and in most cases to ±30%. The set of partition coefficients obtained by Kato et al. [40] represents lower limits only, owing to constraints imposed by the experimental method employed.

may have extended as deep as 400 km. This interpretation has become popular in recent years, e.g. [30–32]. The model has also been extended to the case of extensive or complete melting of the entire mantle e.g. [33–35]. According to this latter scenario, the mantle, of chondritic composition, became thoroughly differentiated, yielding a perovskititic lower mantle en-

Fig. 1. Variation of Sm/Hf, Sc/Sm and Si/Mg ratios as a function of Mg-perovskite fractionation from a chondritic mantle composition, calculated from partition coefficients given in Table 1. The Si/Mg ratio of the present upper mantle is also indicated. Shaded areas indicate uncertainties which would be caused by errors of ±2 in the Sm/Hf and Sc/Sm ratios used in the calculation. From Kato et al. [36].
Walker [35]. The imply, moreover, that the composition of the lower mantle is generally similar to that of the present upper mantle.

Kato et al. [36] also tested an alternative scenario The Earth was postulated to have been born in a totally molten state and the mantle then differentiated as postulated by Agee and Walker [35]; however it was assumed that all traces of this early differentiation were subsequently removed via homogenisation caused by pervasive subsolidus mantle convection. The recent discovery of 4.2 b.y. zircons in Western Australia [37, 38] provides an important constraint on this model. These zircons possess 'primitive' hafnium isotope ratios, i.e. their source regions evolved between 4.55 and 4.2 b.y. in a geochemical environment characterised by an approximately chondritic Hf/Lu ratio [39]. If the reasonable assumption is made that the partition coefficient of Lu is similar to that of Yb, the results of Kato et al [36] as depicted in Fig. 2 show that only a few percent of fractionation of MgSiO₃ perovskite into the lower mantle would cause the Hf/Lu ratio of the upper mantle to change beyond the limits recorded by the ancient zircons. Thus, if the mantle experienced extensive melting and differentiation around 4.5 b.y., accompanied by extensive separation of MgSiO₃ perovskite into the lower mantle, it must have become convectively rehomogenised by 4.2 b.y. This requirement encounters certain problems which are discussed below.

(b) Gravitationally Stable Mantle Stratification

Consider the crystallization of a wholly molten pyrolite mantle, normatively equivalent to ~73% MgSiO₃ perovskite, ~20% (Mg, Fe)O magnesiowüstite and ~7% CaSiO₃ perovskite [41, 42]. This composition contains ~0.4% Na₂O and 0.03% K₂O. The course of differentiation can be modelled on the basis of recent data on melting equilibria at very high pressures and temperatures [43, 36]. The melting point gradient of the mantle is greater than the adiabatic gradient and hence crystallization of the liquidus phase, MgSiO₃ perovskite, proceeds from the base of the mantle upwards until it is joined by (Mg, Fe)O magnesiowüstite. Because of the low abundance of CaSiO₃ and the small but nevertheless significant solubility (0.5–2.0%) of CaO in MgSiO₃ (pv), it seems unlikely that CaSiO₃ (pv) would appear as a liquidus
phase during the crystallization of the lower mantle, which would thus consist almost exclusively of MgSiO$_3$ (pv) + (Mg,Fe)O (mw). The crystal/melt partition data [36] as summarized in Table 1 show that by the time crystallization has progressed upwards to a depth of 650 km, the overlying liquid would have become moderately enriched in Al$_2$O$_3$ (5–8%) and strongly enriched in CaO (10–12%) and Na$_2$O (~1.5%) which are essentially excluded from the crystalline lattice of MgSiO$_3$ (pv). At this depth, crystallization of MgSiO$_3$ (pv) and (Mg,Fe)O (mw) would cease, these phases being replaced by pyrope-rich garnet ± γ, β(Mg,Fe)$_2$SiO$_4$ [44].

Crystallization between 650 and 400 km is controlled by the separation of garnet and of γ and β(Mg,Fe)$_2$SiO$_4$. Calculations based upon partition equilibria between majorite garnet, β(Mg,Fe)$_2$SiO$_4$ and liquid, [36,32], Table 1 , together with mass balance constraint indicate that after crystallization had extended upwards to 400 km, the composition of the magma ocean would be ultrabasic, enriched in CaO (~12–14%) and Na$_2$O (~2.0–2.5%) and low in Al$_2$O$_3$ (~4–5%). These characteristics are imposed by the conditions that the garnet and γ, β(Mg,Fe)$_2$SiO$_4$ phases crystallizing between 650–400 km essentially exclude Na$_2$O and partially exclude CaO, whilst Al$_2$O$_3$ becomes significantly depleted in the melt because of its high partition coefficient in garnet (Table 1).

If the magma ocean above 400 km crystallized without further fractionation, it would yield an assemblage consisting mainly of olivine + omphacite (Ca,Na)(Mg,Fe,Al)Si$_2$O$_6$. Orthopyroxene would be absent because of the high CaO content of the liquid whilst garnet would be suppressed or absent because of high (Na + K)/Al. Continued differentiation of the upper mantle (above 400 km) would be dominated by the separation of Mg-olivine and omphacite. Experimental investigations [45,31] indicate that little fractionation of the jadeite component (NaAlSi$_2$O$_6$) of clinopyroxene would occur between depths of 400–200 km. However, above 200 km, sodium would become enriched in residual liquids, this effect becoming stronger at shallower depths. Potassium would be strongly fractionated into residual liquids throughout the entire differentiation sequence.

At the stage by which crystallization had extended upwards to a depth of 30 km or so, the residual liquid would contain 2–3% K$_2$O and 5–8% Na$_2$O together with appreciable contents of H$_2$O. Since jadeite is unstable at low pressure and elevated temperature, the residual liquid would be expected to crystallize to an assemblage containing alkali felspar and nepheline as major phases. Amphibole, biotite and Fe-rich olivine would probably be the principal subsidiary phases. Because of the low density of this assemblage, it would form a highly buoyant primitive crust. This crust would have contained most of the Earth's inventory of highly incompatible and heat-producing elements (e.g. K, Rb, Ba, Sr, Pb, IREE, LREE, U and Th).

The earlier discussion ignored the possibility that omphacite might be less dense than the melt and float in the upper mantle. This could lead to the formation of a buried magma ocean at depths of 200–300 km or so, possessing geochemical characteristics quite unlike those of Archean mantle or its magmatic products. A buried magma ocean, however, would become gravitationally unstable relative to overlying upper lithosphere, once the latter had cooled. Impacts of large meteorites during the first 500 million years or so would be expected to fracture the lithosphere providing pathways for the upward transport of magma. Once these pathways were established, the buried magma ocean would be rapidly drained and thereby become re-established at the surface. Differentiation of this secondary surficial magma ocean would then proceed as discussed previously.

The stratified structure proposed for the differentiated mantle is illustrated in Figure 3. This structure would be gravitationally stable and highly resistant to rehomogenization via thermally driven convection. The crust would be intrinsically buoyant relative to the upper mantle. Likewise, the upper mantle would be intrinsically stable relative to the underlying region because the omphacite component does not transform to denser phases(s) until depths of 500–600 km [46,47]. Thus, any upper mantle material somehow transported below 400 km would be buoyant because of untransformed omphacite which possesses a much lower density than the garnet and β, γ(Mg,Fe)$_2$SiO$_4$ comprising the region between 400–650 km. Likewise, material of the transition zone between 400–650 km is intrinsically gravitationally stable relative to the perovskite lower mantle. Al-rich garnet from this region requires a higher pressure to transform to perovskite than does the relatively Al-poor composition which is characteristic of the lower mantle. (See [44] and [48] for detailed discussion of this point).

The above analysis has ignored the effects of Fe-Mg fractionation. Because early phases would be Mg-rich,
On the other hand, density calculations based upon experimental measurements of Mg/Fe crystal/liquid K$_D$ values for majorite and $\beta$(Mg,Fe)$_2$SiO$_4$ [32] indicate that Fe/Mg fractionation would be unable to overcome the chemical buoyancy of the upper mantle (olivine + omphacite) as compared to the transition zone (garnet $\pm \beta,\gamma$(Mg,Fe)$_2$SiO$_4$). Likewise, the Fe/Mg fractionation effect is far too small to overcome the intrinsic buoyancy of the primitive crust (see below).

These considerations imply that if the mantle had been completely melted at the stage when accretion of the Earth was nearly completed, it would have differentiated so as to produce a gravitationally stable stratified structure which would have been resistant to rehomogenization by thermally-driven convection. It is therefore very difficult to understand how the highly-fractionated upper mantle became rehomogenized with the lower mantle by 4.2 b.y., as would be required to explain the Hf isotopic composition of early Archaean zircons, as discussed previously.

(c) The Primitive Crust

We concluded above that crystallization of the entire mantle from a molten state would ultimately produce a buoyant primitive crust with alkali felspar and nepheline as major phases. This crust may have been about 30 km thick and would have contained most of the Earth’s inventory of highly incompatible and heat-producing elements. The composition of this primitive crust would have been dramatically different from that of the continental crust which is known to have existed at 3.8 b.y. All traces of the existence of this primitive crust must therefore have been removed prior to this time. It is difficult to understand how this could have occurred. Even if plate tectonics were operative at this very early stage, subduction of highly buoyant primitive crust appears unlikely. If indeed it occurred, subducted primitive crust would fuse at shallow depths because of its low melting point and the high temperature of surrounding mantle. Most of the melted material would have been promptly returned to the surface. Extraction of most of the radioactive heat sources from the mantle into the crust would diminish the driving forces for thermal convection in the mantle. Once a 30 km thick, buoyant, low-melting point crust containing most of the Earth’s K, U and Th had been formed, it would be unlikely ever to be destroyed by subsequent convection.
3. Core Formation in Planetary Bodies

More than a decade ago, Heinrich Wänke and his colleagues in Mainz developed a powerful set of geochemical arguments pointing towards a genetic relationship between the Moon and the Earth’s mantle [49–52]. Independently, in Canberra, my colleagues and I reached similar conclusions [53–56]. Both groups recognised certain key resemblances between the siderophile element abundances in the Earth’s mantle and in the Moon, and concluded that most of the material now in the Moon had been derived ultimately from the Earth’s mantle. This conclusion generated intense controversy at the time. Most lunar scientists were reluctant to accept the proposition that the vexed question of the Moon’s origin could be settled by such a “simple” argument. Nevertheless, the controversy had beneficial results because it stimulated the acquisition of additional analytical data on the abundances of siderophile elements in lunar and terrestrial rocks and on the partition coefficients of siderophile elements between silicate and metal phases.

I believe that the principal geochemical issues relating to this controversy have been clarified during the intervening years. Nevertheless it remains doubtful whether the implications of recent developments in siderophile geochemistry have been realized by many lunar scientists.

When a metallic core forms within a planetary body, varying proportions of siderophile elements are left behind in the silicate mantle. The resultant siderophile signatures in the mantles reflect the nature of the core-forming process in the particular planet. This may be illustrated by reference to the core forming processes in the Earth, Mars and the eucrite parent body (EPB).

The abundance patterns displayed by siderophile elements in the Earth’s mantle are extraordinarily complex. A simplified account of the major trends may be summarized as follows [57–60]:

- A major group of siderophiles including Ni, Co, Cu, Fe, Ga, W, Mo, P, As, Sb and Ge are present in abundances ranging from 1–15 percent of their abundances in CI meteorites.
- Several of these elements (eg. Ni, Co, Mo) are much more abundant in the mantle than would be expected if they had been partitioned into the Earth’s core on the basis of their iron/silicate partition coefficients as determined under low pressure conditions
- The relative abundances of these elements in the mantle are not significantly correlated with the above metal/silicate partition coefficients.
- The platinum-group metals and also gold and rhenium display quite different abundance characteristics to this first group. In contrast to the highly variable abundances (normalized to CI chondrites) of the first group, the ‘noble’ metals are present in chondritic relative abundances, but at levels corresponding to about 0.3% of CI abundances. These levels are also far higher than would be expected on the basis of low-pressure iron/silicate partition coefficients.
- Vanadium, chromium and manganese are depleted in the mantle, as compared to CI abundances (Mg-normalized) by factors of 0.6–0.2. Although these elements are not normally regarded as being “siderophile”, increasing evidence (discussed later) suggests that the depletions may be due at least in part to their incorporation in the Earth’s core.

This highly complex signature of the siderophile abundance patterns in the mantle is the end product of several physico-chemical processes which were involved in the formation of the Earth’s metallic core. The following factors were among those involved:

- Partitions according to chemical affinities of siderophiles between a predominantly metallic iron phase amounting to ~30% of the mass of the Earth and mantle silicate phases during the formation of the Earth’s core.
- The circumstance that these partitions occurred within a body of planetary dimensions under conditions of very high pressure and temperature. Metal/silicate partition coefficients for siderophile elements may have been substantially influenced by these high P, T conditions.
- The processes which formed the Earth’s core also led to the incorporation of about 10 (wt) percent of light elements. It is now believed that oxygen is the principal member of this group [61, 58, 62]. The presence of the light element component would necessarily influence the partitions of siderophile elements between core and mantle.
- It is unlikely that chemical equilibrium between metal and silicates was achieved during the formation of the Earth’s core [24, 25, 3]. Moreover, several widely supported cosmogonic hypotheses have
proposed that the Earth accreted inhomogeneously e.g. [63, 64] in a manner which would have prevented chemical equilibrium being achieved between core and mantle.

The detailed operation and relative importance of the several physico-chemical processes involved in formation of the Earth’s core are poorly understood. However, it is readily recognized that this particular combination of complex processes is most unlikely to have operated similarly in all planetary bodies. It seems reasonable to assume, therefore, that the siderophile signature of the Earth’s mantle should be unique to the Earth, or, at the very least, to a planet of similar size which had experienced similar geo-chemical evolution. Venus is the only candidate which could be in this class.

The relative abundances of siderophile elements in the mantle of Mars and of the eucrite parent body (EPB) have been estimated by Dreibus and Wänke [65, 66] and are shown in Figure 4. It is seen that they are quite different from the Earth’s mantle and also differ markedly from each other. These differences add credence to the view expressed above that the mantles of differentiated planetary bodies are characterized by unique siderophile signatures.

The mantle of Mars is depleted in Cu, Ni, Co and As (Fig. 4) as compared to the Earth’s mantle. In contrast, the martian mantle is enriched in P, W, Mn and Cr which appear to be present in near-chondritic abundances (Mg-normalized). Dreibus and Wänke [65] and Treiman et al. [67] have pointed out that these characteristics could be explained if the martian core were rich in sulphur. The mantle of the eucrite parent body is depleted in Ni, Co, Cu, Mo, Ga, Ge, Ir as compared to the Earth’s mantle and enriched in Mn, V and Cr. Phosphorus and tungsten are present at levels similar to those in the Earth’s mantle but are depleted as compared to the martian mantle. These characteristics are readily explained if the EPB core had been composed dominantly of metallic iron and was depleted in sulphur relative to the martian core [55, 56, 68].

4. The Lunar Siderophile Signature

Geophysical and geochemical data strongly suggest that the Moon possesses a small metallic core, probably amounting to about 1–2 percent of the lunar mass, e.g. [69, 72]. Experimental investigations show that if the lunar core had formed by segregation of metal initially dispersed throughout the lunar mantle, it would have contained about 40 percent of nickel [73, 74]. This tiny core formed under vastly different pressures, temperatures and other physico-chemical conditions to those prevailing during formation of the Earth’s core. In line with the previous discussion, it would be expected that the lunar siderophile signature would differ drastically from that of the Earth’s mantle. In fact, as the following examples show, there are some remarkable similarities in the siderophile abundance pattern of the Earth’s mantle and Moon. There are also some important differences, but it will be seen that these are readily explained by the operation of well-known processes.

(a) Tungsten

Correlations between W and U in lunar and terrestrial rocks, based upon the work of Wänke and colleagues [75–77] are shown in Figure 5. In oxidized form, W is an incompatible element and behaves similarly to U during magmatic fractionation. Hence, the W/U ratio in terrestrial rocks remains approximately constant over a wide concentration range (Figure 5). Tungsten and uranium are both refractory elements which were accreted by the Earth in chondritic relative abundances. However, W became depleted in the Earth’s mantle by a factor of about 25 owing to its siderophile nature which caused it to enter the core. It
(b) Phosphorus

The behaviour of phosphorus in the Earth-Moon system is analogous to that of tungsten (Figure 6). Phosphorus displays a similar degree of incompatibility to neodymium in terrestrial fractionation processes [78]. However, phosphorus is also siderophile and has been depleted by a factor of about 20 in the Earth's mantle owing to its entry into the core. It is seen in Fig. 6 that phosphorus has been depleted by a factor of about 30 in primitive mare basalts [54] and in KREEP and lunar highland rocks [49, 55, 56]. This depletion could not have been caused by extraction into the tiny lunar core. (The small difference between terrestrial and lunar P abundances may have been caused by selective loss of P in a volatile species from protolunar material, e.g. [69]).

(c) Cobalt

Unlike W and P, Co is a compatible element which is not strongly fractionated (relative to Mg and Fe) during magmatic differentiation. Abundances of Co in terrestrial tholeiites and lunar low-Ti basalts are shown in Fig. 7 and are seen to be very similar, implying similar abundances in their respective source regions. The abundances of indigenous cobalt (designated Co*) in Apollo 16 highland breccias are plotted against total (Mg + Fe) in Figure 8. Indigenous Co* contents of these rocks were obtained by subtracting meteoritic cobalt contamination using the chondritic Co/Ir ratio and the assumption that all of the iridium in the highlands was derived from meteoritic contamination [55, 56]. Figure 8 shows that there is a clear correlation between residual Co* and (Mg + Fe) for Apollo 16 highland breccias. Moreover, the line of best fit extrapolates directly to the least fractionated volcanic glasses from the lunar mantle, showing that the correlation is of global significance. Figures 7 and 8 imply that the bulk cobalt content of the lunar mantle is similar to that of the Earth's mantle. Ringwood and Seifert [73] have provided a more detailed discussion of lunar and terrestrial cobalt geochemistry.

(d) Nickel

The relationships between nickel contents of primitive lunar volcanic glasses [79] and mare basalts, and their respective MgO/(MgO + FeO) ratios is shown in Figure 9. The corresponding field for terrestrial ocean floor tholeiites is also shown. It is seen that the
two fields almost completely overlap. This can hardly be a coincidence. It implies that lunar and terrestrial mantle magmatic systems possessed similar amounts of nickel at corresponding stages of differentiation as indicated by Mg-numbers.

However, the situation is not quite as straightforward as this because the absolute abundance of FeO in the lunar mantle is believed to be substantially higher than in the terrestrial mantle, e.g. [81]. Seifert et al. [74] demonstrated that the nickel content of the lunar mantle source region which produced the most primitive lunar volcanic glass yet recovered contained about one third of the nickel abundance of the terrestrial mantle. In view of the very limited sampling of lunar surface materials it is possible that still more primitive volcanic glasses possessing higher nickel contents are present, implying correspondingly higher nickel contents in their source regions.

This conservative lower limit of ~ 830 ppm Ni in olivine from the lunar mantle, [71, 74] still implies that the lunar mantle resembles the Earth’s mantle in this respect much more than it does the mantles of the eucrite parent body and Mars which are highly depleted in nickel. Delano and Ringwood [55] and Seifert et al. [74] showed that the metal phase in equilibrium with the most primitive mare basalts and volcanic glasses contains 35–40 percent of nickel. The lunar core is thus expected to be nickel-rich. The lunar
Fig. 8. Cobalt in primitive lunar volcanic glasses and Apollo 16 breccias and soils (corrected for meteoritic contamination) versus (Mg + Fe). Note that the solid line representing the best fit to lunar highland data projects directly into the field of primitive volcanic glasses showing that the correlation is of global significance. After [80].

Fig. 9. Nickel contents versus MgO/(MgO + FeO) ratios in primitive lunar volcanic glasses [79] and Apollo 12 basalts compared with corresponding field observed for terrestrial basalts. The boundaries of the latter are shown by the stippled region inside the solid lines. Based on [56, 73].

The nickel abundance systematics of lunar highlands rocks are complicated by widespread meteoritic contamination. Wänke et al. [52] and Delano and Ringwood [55, 56] assumed that the contaminating meteoritic material was of chondritic composition. Making the conservative assumption that all of the iridium in the lunar highlands was supplied by meteoritic contamination, they subtracted out meteoritic nickel contamination using the chondritic Ni/Ir ratio, and found substantial nickel residuals in the highland breccias which were believed to represent indigenous lunar highland nickel. These residuals (designated Ni*) were quite well correlated with Mg (Fig. 10), with total (Mg + Fe), and with indigenous cobalt (Co*) in lunar breccias [52, 55, 56, 71, 80]. The existence of these correlations provided strong evidence that the residual nickel (Ni*) was indeed indigenous.
Delano and Ringwood [55, 56] and Ringwood et al. [80] also observed systematic relationships in lunar highland breccias and soils between Ni* and Co*. Plots of Ni* versus Ni*/Co* for lunar highlands samples are compared with corresponding plots for total Ni and Co in terrestrial basalts and komatiites in Figs. 11 and 12. It is seen that the Ni* versus Ni*/Co* plots display remarkably systematic relationships in lunar highlands samples. Moreover, they are very similar to corresponding plots for terrestrial oceanic tholeiites. At higher Ni* levels, as in Apollo 16 breccias (Fig. 12), the lunar trend diverges somewhat from that of terrestrial komatiites but the overall resemblance in the two trends is nevertheless strong.

The Ni/Co versus Ni trend in terrestrial basalts and komatiites arises from variable degrees of partial melting and fractional crystallization involving olivine-melt equilibria. Ringwood et al. [80] suggested that the same basic process is responsible for the lunar trends. They also pointed to several factors which could be responsible for the divergence between lunar and terrestrial komatiite trends at high nickel levels (Figure 12). Overall, the data shown in Figs. 11 and 12...
clearly demonstrate the existence of “earthlike” Ni* and Co* abundance systematics in lunar highland breccias.

Despite the compelling evidence that the lunar mantle contains substantial amounts of indigenous nickel and cobalt, and the general acceptance that a large proportion of the cobalt in lunar breccias was indigenous, the proposal by Wänke, Delano and Ringwood that a substantial part of the nickel in the lunar highlands was also indigenous aroused great consternation among lunar scientists. The prevailing view was that essentially all of the nickel in the lunar highlands had been provided by contamination from meteorites possessing non-chondritic Ni/Ir, Ni/Co and Ni/Au ratios. Unfortunately, this combination of non-chondritic ratios has not been found in any known class of meteorite. The meteoritic material which bombarded the Earth during the first 500 m.y. or so and became mixed into the mantle was dominantly chondritic, based upon platinoid-Re-S-Se-Te systematics [82]. This conclusion stands in contrast to the arbitrary assumption that large areas of the lunar highlands were heavily contaminated by an unknown class of meteoritic projectiles.

The controversy seems to have been sustained by the belief held by many lunar petrologists that the key to understanding the nature of the lunar crust was provided by the so-called “pristine” lunar rocks which are minor components of the highland rock suite. Most, (but not all, eg. 72417) pristine rocks possess low abundances of indigenous nickel. However Wänke et al. [51] pointed out that most pristine rocks are derived from magma bodies which have crystallized relatively slowly at high levels in the lunar crust under low oxygen fugacities. Under these conditions, auto-reduction is likely to occur, leading to the precipitation of metallic iron, which would extract nickel and other siderophiles. Thus the residual liquids would acquire “artificially” low siderophile element contents. Precisely the same effect is observed during the crystallization of mare basalts at the lunar surface. Low nickel contents in pristine rocks do not necessarily reflect the nickel abundances of their source regions, and their relevance to the indigenous nickel content of highland rocks is dubious, to say the least a.

The significance of the relationships depicted in Figs. 7–12 should be readily apparent. They testify to a close relationship between the geochemistry of cobalt and nickel in the Earth’s mantle and in the Moon. This testimony is clearly of profound genetic importance.

---

a The geochemistry of siderophile elements in the lunar highlands has recently been discussed by Warren et al. [161]. These authors maintain that the compositions of lunar highland rocks (and particularly A 16 samples) retrieved during the Apollo program were atypical and that the compositions of six lunar meteorites believed to be derived from the lunar farside provide a better representation of the average lunar highlands composition. They concede that A 16 rocks may contain a significant proportion of indigenous Ni but believe that this is not typical of the lunar highlands generally. Irrespective of the merits of their arguments, they have seriously misunderstood a key aspect of Ringwood et al.'s. interpretation of lunar highland geochemistry [80]. The latter authors identified a primitive komatiitic component, rich in Ni, only at the A 16 site. Warren et al. [161] quoted us as proposing that this primitive Ni-rich komatiitic component is characteristic of lunar highland breccias generally. This is entirely incorrect. We concluded that the primitive components of highland breccias at the A 14, 15 and 17 sites contained much lower levels of indigenous nickel and suggested that the primitive components at these sites were derived from internally generated basaltic magmas rather than komatiites. More generally, Ringwood et al. [80] stated: “The highland breccias do not contain a single primitive component. Instead, this component is believed to be represented by a wide range of petrologic compositions varying from basic to ultrabasic”. The data presented by Warren et al. [161] suggest that the primitive components at the sites represented by the lunar meteorites may also have been basaltic, rather than komatiitic.
(e) Chromium, Vanadium and Manganese

The abundances of these elements (Mg-normalized) are very similar in the Moon and Earth’s mantle e.g. [51, 80, 83, 84, 85]. However, they are substantially depleted in the mantles of both bodies as compared to Cl chondrites. The depletion factors (references above) are V: 0.6–0.7; Cr: 0.4–0.5; Mn: 0.25–0.35. Dreibus and Wänke [65, 66] showed that Mn, Cr and V are present in near-chondritic abundances in the mantle of the EPB and that Mn and Cr were likewise undepleted in the martian mantle [70]. (The martian vanadium abundance is poorly constrained).

Ringwood [25] suggested that the depletions of Mn, Cr and V in the Earth’s mantle were caused by their entry into the core. Wänke [64] proposed that the Earth accreted inhomogeneously and that V, Cr and Mn became siderophile under the highly reducing conditions which prevailed during the early and intermediate stages of accretion of the Earth. Dreibus and Wänke [65] also proposed an alternative explanation, suggesting that the extensive solubility of FeO in molten iron at high pressures and temperatures inferred by Ringwood [61] and later confirmed [62] might enhance the partitions of Mn, Cr, V into the metallic phase.

Drake et al. [85] proposed that the depletions of Mn, Cr and V in Earth and Moon were caused by volatility-controlled fractionations in the solar nebula prior to accretion. This explanation is untenable for V which condenses at a temperature about 110 K higher than Mg. It is also implausible in the case of Cr which condenses only 35 K below Mg [86]. The Mg/Cr ratio is near-constant in different chondrite groups (Cl = 36, CM = 38, CO = 41, CV = 41, OL = 40, OH = 42, E(average) = 41) and is uncorrelated with Si/Al ratios in these groups which range from 9 to 20 and have probably been fractionated by processes involving selective volatility [65, 76].

The nebula condensation temperature of Mn is ~150 K below that of Mg and it is possible that the Mn depletions in the Earth and Moon were caused by its higher volatility. However, this interpretation is not supported by the observation that Mn is undepleted in the mantle of the eucrite parent body which experienced much stronger depletions of many other volatile elements than did the Earth’s mantle [65]. Moreover, Mn is slightly more abundant in the lunar mantle than the terrestrial mantle [54] despite much greater depletion of most other volatiles in the Moon.

We conclude that the depletions of V, Cr and possibly Mn in the Earth have most probably been caused by their siderophile behaviour under the conditions of core formation in a large planet. Kato and Ringwood [87], and Drake et al. [85] showed experimentally that core formation processes in martian-sized and smaller planetary bodies would not deplete V, Mn and Cr in the mantles of these bodies. This is entirely consistent with the observation that the mantles of the EPB and Mars are not depleted in these elements.

The similarity of abundances of V, Cr and Mn in the lunar and terrestrial mantles assumes great significance in the light of these circumstances, as was first pointed out by Dreibus and Wänke [65].

(f) Highly Siderophile Elements

Elements which are more strongly siderophile than nickel are depleted in lunar basalts as compared to terrestrial basalts by quite large factors varying from 10–100. These depletions are probably characteristic of their source regions in the lunar mantle. The depleted elements include rhenium, gold and molybdenum [54, 88, 89]. Iron-silicate partition coefficients for these elements range from $10^3$–$10^6$ and accordingly, their abundances are sensitively influenced by small degrees of metal fractionation. Wänke et al. [90] and Newsom [69, 89] have shown that the depletions could be readily explained by the separation of a small metallic core within the Moon amounting to as little as one percent of its mass. Independent geophysical evidence supports the existence of a very small metallic core within the Moon [72, 91].

(g) Volatile Siderophile Elements

It is well known that the Moon has experienced strong depletions of many volatile elements as compared to the Earth’s mantle. The depleted species include both lithophile elements (eg. Na, K, Rb, F, Zn, Cl) and siderophile elements (eg. Ge, Sn, Cu, As, Ga). The abundances of these volatile siderophile elements in protolunar material prior to the volatilization event which caused their depletions were estimated by Seifert and Ringwood [92] on the basis of comparative abundances in terrestrial and lunar low-Ti mare basalts. It was assumed that the various siderophiles had been depleted by volatilization to a similar extent as had lithophile elements possessing similar volatilities, i.e., similar nebula condensation temperatures. After correction for volatilization losses in this man-
The abundances of Ga, Sn and Cu in protolunar material were found to be similar (within a factor of two) to those in the Earth's mantle. These elements possess metal/silicate partition coefficients between 4 and 100. On the other hand, estimates of Ge and As abundances in protolunar material remained smaller than the terrestrial mantle abundances, even after volatility corrections had been applied. However, these two elements possess metal-silicate partition coefficients greater than 1000 and hence it seems plausible that their depletions also reflect their extraction into the lunar core, in a manner similar to Mo, Re and Au.

Sulphur and selenium are two moderately siderophile elements which possess similar abundances in low-Ti lunar basalts and terrestrial oceanic tholeiites [54]. These elements are usually regarded as being highly volatile in cosmochemical processes, where they react readily with hydrogen to form H₂S. However, under conditions of low hydrogen fugacity, in environments where H₂/H₂S and H₂/H₂O ratios are much smaller than in the solar nebula, S and Se can be relatively involatile, particularly when dissolved in silicate melts as S⁻² ions. These conditions may have been relevant to those under which the Moon was formed [54].

(h) Implications of Lunar Siderophile Signature

Ratios of the abundances of siderophile elements in the lunar and terrestrial mantles are plotted against their metal (iron)-silicate partition coefficients in Figure 13. In the cases of volatile siderophiles (Cu, Sn, Ga, Ge, As), the abundances have been corrected to correspond to the levels in protolunar material, prior to partial loss by volatilization, as described by [92]. Diagrams analogous to, but less comprehensive than Fig. 13 have been constructed by Wänke and Dreibus [76] and by Newsom [69]. It is seen that the abundances of elements less siderophile than Ni (Mn, V, Cr, Ga*, Fe, Sn*, Cu*, W, P, Co, S, Se) are similar (within a factor of 2) in both bodies. Nickel itself is depleted by a factor ~ 3 in the source regions of the most
pristine lunar basalts but is present in terrestrial abundances (relative to Mg) in breccias at the A 16 lunar highlands site.

The similarities in abundances of a large group of siderophiles in terrestrial and lunar material is quite remarkable. It was pointed out previously that because of the complexity and multiplicity of processes connected with core-formation, which ultimately established the siderophile pattern of the terrestrial mantle, this pattern was likely to be unique to the Earth. As seen in Fig. 4, it is quite different from those in the EPB and Mars. No way is known whereby such a pattern can be established in the solar nebula prior to planetary accretion, nor by internal differentiation of a core within a lunar-sized body, nor even by internal differentiation of a core within a martian-sized body [87]. The similarities in terrestrial and lunar abundances of this group of siderophiles provide powerful support for the hypothesis that a large proportion of the material now in the Moon was derived from the Earth’s mantle subsequent to core-formation.

It is seen from Fig. 13 that elements more siderophile than Ni (Ge*, Mo, As*, Re, Au) are depleted in the lunar mantle as compared to the terrestrial mantle. The degrees of depletion of this group of elements correlate roughly with their metal-silicate partition coefficients [69]. These depletions can readily be explained by separation of a small metallic core amounting to <1% of the lunar mass [76, 69]. This would not markedly affect the abundances of moderately siderophile elements but would have strongly depleted the highly siderophile elements. The presence of a small (<2% by mass) lunar core is strongly suggested by the lunar moment of inertia coefficient [94] and by the observation of a phase shift in the forced precession of the lunar figure [91].

(i) Isotopic Relationships

The isotopic compositions of lunar strontium and xenon also suggest a close relationship to the Earth. Carlson and Lugmair [95, 96] obtained a crystallization age of 4.44 ± 0.02 b.y. for a pristine lunar anorthosite. Its strontium isotopic composition at that time was surprisingly radiogenic and suggested that it had evolved in an environment possessing a terrestrial Rb/Sr ratio in the interval between 4.55 and 4.44 b.y. and quite unlike the present lunar Rb/Sr ratio. A drastic depletion of Rb must have occurred near 4.44 b.y., possibly associated with formation of the Moon from the Earth, or with loss of Rb from the lunar magma ocean from which the anorthosite crystallized.

Some lunar highland rocks have trapped a xenon component possessing terrestrial isotopic composition, quite unlike that of other major xenon reservoirs in the solar system [97, 98]. There was some ambiguity about the significance of this result because of the possibility of adsorption of terrestrial xenon during fine grinding of the sample [99]. However, xenon in lunar sample 60025 consisting of unground chips, also possessed a terrestrial signature [97]. This anorthosite also contained anomalously high amounts of other volatile elements that are certainly of lunar origin [100].

(j) Independent Planet Hypothesis

The siderophile patterns shown in Figs. 4 and 13 introduce formidable difficulties for hypotheses which maintain that the Moon formed as an independent planet which, owing to special circumstances, had accreted in orbit around the Earth or had been captured into Earth orbit. Newsom and Drake [101, 102] have attempted to show that the lunar and terrestrial abundances of individual siderophile elements can be explained with the aid of appropriate ad hoc assumptions. However, attempts to explain simultaneously the common abundances of more than 2 or 3 pairs of siderophile elements on the basis of a self-consistent and physically realistic set of assumptions have been unsuccessful.

Independent planet hypotheses encounter the first order problem of explaining the drastic depletion of metal phase in the Moon as compared to Earth, Mars and Venus. Previous attempts to explain this problem which rely on some kind of mechanical fractionation of silicates from metal seem contrived. Several workers have previously proposed that accretion of terrestrial planets was preceded by the formation of an earlier generation of asteroidal-sized planetesimals that melted and differentiated to form metallic cores overlain by silicate mantles. The differentiation is believed to be analogous to that experienced by the parent bodies of eucrites, pallasites and iron meteorites. Chapman and Greenberg [103] speculate that collisional processes occurring in a circumterrestrial swarm of planetesimals permitted the Earth to selectively capture the metallic cores of these bodies.
whereas the disintegrated silicate mantles were captured into the circumterrestrial swarm that subsequently coagulated to form the Moon. Related models have been proposed by other workers [12, 104–106].

Models of this type predict that the chemistry of the Moon should be similar to that of the eucrite parent body and also imply that the silicates in the Moon had equilibrated with iron-rich metal at low pressures and at temperatures less than 1300 °C before being filtered out by the circumterrestrial ring. The observed composition of the Moon, and in particular, the high Co and Ni content of the lunar mantle and the depletions of Cr, V and Mn squarely contradict these hypotheses. Further detailed discussion of the problems encountered by the independent planet hypothesis in the light of the siderophile evidence is given in [59, 73, 93].

5. Significance of the Lunar Bulk Composition

(a) Compositional Models

Estimates of the bulk composition of the Moon are necessarily model-dependent. Three such estimates are given in Table 2. The first is based upon the compositions of 25 defined classes of pristine lunar volcanic glasses representing the most primitive mafic and ultramafic magmas to have been erupted from the lunar interior [107]. The composition of their source regions in the lunar mantle is derived from compositional relationships among the glasses together with known chemical equilibria between melt compositions and refractory residual phases obtained from experimental petrology. The methodology is analogous to that by which the composition of the Earth’s mantle is obtained from primitive terrestrial basalts and refractory peridotitic residua [29].

The second estimate of lunar bulk composition is obtained from the chemistry of lunar highland breccias. Wänke et al. [83, 108, 111] showed that their compositions could be interpreted in terms of mixtures between two main components – (i) anorthosite and (ii) a “primary” component rich in Mg and Cr and also containing substantial indigenous Ni and Co (Figure 14). Wänke and colleagues obtained the composition of the primary component by extrapolating the mixing lines of Fig. 14 to a point where Ca and Al displayed the chondritic ratio. Ringwood et al. [80] showed that the composition of the primary component present in Apollo 16 highland breccias was similar to that of a komatiite (Table 3). They proposed that after the differentiation of the primitive lunar magma ocean which produced an anorthositic crust, extensive partial melting occurred at greater depths in the Moon, producing komatiitic magmas which were erupted at the lunar surface and became intimately mixed with the anorthositic component via meteoritic impact processes. A bulk composition for the Moon was derived from the compositions of komatiitic and residual olivine, analogous to the procedures used for obtaining the composition of the Earth’s mantle from terrestrial komatiites [41]. The composition of the lunar

### Table 2. Comparison of lunar and terrestrial mantle compositions.

<table>
<thead>
<tr>
<th></th>
<th>Lunar</th>
<th>Terrestrial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>SiO₂</td>
<td>42.6</td>
<td>43.2</td>
</tr>
<tr>
<td>TiO₂</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>Cr₂O₃</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>FeO</td>
<td>13.6</td>
<td>12.2</td>
</tr>
<tr>
<td>MnO</td>
<td>0.2</td>
<td>0.16</td>
</tr>
<tr>
<td>MgO</td>
<td>37.1</td>
<td>36.9</td>
</tr>
<tr>
<td>CaO</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.06</td>
<td>0.09</td>
</tr>
<tr>
<td>Ni ppm</td>
<td></td>
<td>2437</td>
</tr>
<tr>
<td>Co</td>
<td></td>
<td>95</td>
</tr>
<tr>
<td>V</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Sc</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

A: Jones and Delano [107]. B: Ringwood et al. [80] terrestrial MgO/SiO₂ ratio. C: Taylor [109]. D: Sun [41]. E: Hart and Zindler [110].

### Table 3. Comparison of lunar and terrestrial komatiites.

<table>
<thead>
<tr>
<th></th>
<th>Lunar</th>
<th>Terrestrial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>SiO₂</td>
<td>45.2</td>
<td>45.2</td>
</tr>
<tr>
<td>TiO₂</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>6.6</td>
<td>6.2</td>
</tr>
<tr>
<td>Cr₂O₃</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>FeO</td>
<td>14.8</td>
<td>11.5</td>
</tr>
<tr>
<td>MnO</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>MgO</td>
<td>26.9</td>
<td>28.6</td>
</tr>
<tr>
<td>CaO</td>
<td>5.3</td>
<td>5.9</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Ni ppm</td>
<td>1327</td>
<td>1396</td>
</tr>
<tr>
<td>Co</td>
<td>78</td>
<td>107</td>
</tr>
<tr>
<td>V</td>
<td>112</td>
<td>124</td>
</tr>
<tr>
<td>Sc</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

A: Primary component of lunar highlands derived from Fig. 14 from Wänke et al. [108] and Ringwood et al. [80]. B, C, D: Terrestrial komatiites from Ringwood et al. [80].
Fig. 14. Binary mixing diagram for highland breccias mostly from the Apollo 16 site. The compositions can be readily interpreted in terms of mixing between an anorthositic end-member and a "primary" component defined by the chondritic Ca/Al ratio of 1.08. Solid lines were fitted to compositions of KREEP-poor samples represented by filled symbols. Samples with high proportions of KREEP are represented by open symbols and were not used in computation of the element mixing lines. (From [108]. Fig. 1).

The lunar mantle so derived is seen to be very similar to that obtained from the pristine volcanic glasses (Table 2).

A third estimate of lunar bulk composition (Table 2) is obtained primarily from aluminium mass-balance considerations between the lunar crust and mantle [109]. This composition depends sensitively upon estimates of the total Al content of the lunar crust, the proportion of the Moon which differentiated to form the crust and the amount of Al which remained behind in the refractory, residual region. These estimates are rather subjective. The lunar bulk composition thereby derived by Taylor [109] is about 50 percent richer in Al₂O₃, CaO and other refractory oxides than compositions based directly upon the petrogenesis of lunar volcanic and highland rocks. The Al mass balance constraints are quite permissive and cannot be used to exclude the less refractory bulk compositions.

It was once thought that observations of lunar heat flow favoured the refractory-enriched composition. However, further geophysical analysis showed the data could also be reconciled readily with the less-refractory bulk compositions [81, 112]. Recently, it has been suggested that refractory-enriched bulk compositions can be reconciled more readily with the distributions of seismic velocities in the lunar mantle and in particular, with the presence of a seismic discontinuity near 500 km [72, 113]. However, the seismic velocity gradients and existence of a discontinuity near 500 km are sensitively dependent upon the assumption that there is no significant degree of seismic velocity anisotropy in the lunar upper mantle. Since this region is believed to be olivine-rich and is likely to have experienced convection at some stage of its history, it would be surprising if a significant degree of velocity-anisotropy were not present. Accordingly, it should be recognized that the presently preferred seismic depth-velocity profiles [114] provide only a weak constraint on the refractory content of the lunar mantle. On the other hand, the mean seismic velocities between the surface and depths of about 1000 km are quite well known and do provide significant compositional and mineralogical constraints. These mean velocities are
entirely consistent with the low-refractory bulk compositions A and B of Table 2.

(b) Earth-Moon Comparisons

Lunar compositional models are also compared to the estimated composition of the terrestrial upper mantle in Table 1. It is seen that Models A and B possess compositions which are similar to the terrestrial mantle, with the important exceptions that they are enriched in FeO and depleted in Na₂O (and other volatiles). Lunar Model C departs markedly from the terrestrial composition. An important observational constraint is that the isotopic composition of oxygen, which amounts to ~ 50% of the Earth’s mantle and Moon, is identical in both bodies [115].

Much has been made in the literature concerning the significance of differences between the bulk compositions of the Moon and Earth’s mantle for hypotheses of lunar origin. In literally dozens of papers, it has been asserted that the existence of significant chemical differences, real or imagined, between both bodies constitutes significant evidence that the Moon could not have been derived primarily from the Earth’s mantle. This is a highly simplistic view. It is only the most extreme (and implausible) versions of the “fission” hypotheses that imply identity in bulk compositions. The models for a terrestrial origin of the Moon which I have developed over many years [25, 26, 59, 81, 93, 116, 117], all lead to significant differences between the bulk compositions of Earth’s mantle and Moon. The same applies to recent models of impact-induced earth-fission, e.g. [76].

Differences between the compositions of Earth’s upper mantle and Moon in the context of models for a terrestrial origin of the Moon may arise from the following factors:

(i) Chemical differences between the two bodies may arise because of fractionation caused by volatilization-condensation processes during removal of protolunar material from Earth and reassembly in Earth-orbit [26, 117].

(ii) The Moon may now incorporate a significant proportion of an “impactor” compositionally different from the Earth’s mantle, which was responsible for ejection of protolunar material from the Earth [76, 118].

(iii) The Earth’s core is believed to contain a large amount of FeO which was extracted from the mantle during the core-formation process [58, 61, 62, 81]. Removal of protolunar material from the Earth’s mantle may have occurred before the FeO-extraction process had been completed [119, 120].

(iv) Ejection of protolunar material may have occurred via multiple impacts of large planetesimals into a chemically differentiated, ~ 200 km deep terrestrial magma ocean [81]. (This could explain enriched Ca and Al abundances in the Moon should this be required by future evidence.)

(v) The Moon is believed to have accreted from a disk of Earth-orbiting planetesimals which themselves were formed from material ejected from Earth’s mantle, e.g. [121]. This protolunar disk may have captured a significant proportion of planetesimals directly from the solar nebula [26].

(vi) Formation of a small metallic lunar core by internal differentiation within the Moon may account for differences in abundances of highly siderophile elements between the lunar and terrestrial mantles as discusses earlier.

(c) Discussion

The factors considered in the previous section emphasize the need for caution in drawing conclusions about lunar origin from simple comparisons of bulk compositions. The same factors would also influence comparisons of siderophile signatures in both bodies. I have never maintained that these patterns are identical. It is the similarity of a large group of siderophile element abundances in both bodies, within a factor of about two, which is of genetic significance for reasons advanced previously. This unique compositional signature has survived despite the effects of other factors, mentioned above, which would tend to degrade it.

Taylor [109] argued that differences in composition between Earth’s mantle and Moon were caused not by the factors discussed in the previous section, but by the circumstance that the protolunar material was derived predominantly from the mantle of a martian-sized impactor which collided with the Earth. He states:

“A seductive feature of the (giant impact) model is that it allows for an ad hoc composition for the impactor within the general constraints imposed by inner planetary compositions. Accordingly, appropriate Fe/Mg ratios, siderophile element contents, and Al
6. Formation of the Moon

(a) Historical Perspective

It is well known that the densities of terrestrial planets vary substantially, even when corrected to a common pressure to allow for the effects of differential self-compression in their own gravitational fields [122], Urey [14], and many others explained the density variation by assuming that these planets are composed of varying proportions of silicate ($\rho_0 = 3.3 \text{ g/cm}^3$) and nickel-iron ($\rho_0 = 7.9 \text{ g/cm}^3$) phases, each phase being essentially of constant composition. According to their model, Mercury and Mars contain about 65 and 20 wt% of metal phase respectively, while Earth and Venus possess intermediate metal contents. On the basis of this hypotheses, the Moon (density 3.334 g/cm$^3$) would contain less than 5% of metal phase.

Urey [14, 123, 124] proposed that physical processes occurring in the solar nebula prior to accretion of the planets had caused variable degrees of fractionation of iron particles from silicate particles in different regions of the nebula. Subsequent accretion of planets in particular regions reflected these pre-existing metal-silicate inhomogeneities. Some kind of metal-silicate fractionation is clearly necessary to explain the high density of Mercury and the low density of the Moon as compared to Mars, Earth and Venus. Moreover, the major role attributed to metal-silicate fractionation in the solar nebula was supported during the 1960s by measurements that seemed to show that the abundance of iron in the sun (relative to silicon) was a factor of five lower than is found in Earth, Mars, Venus, and chondritic meteorites. Urey's interpretation accordingly became widely accepted and was reflected in many other subsequent cosmogonic models, e.g. [125–127].

Ringwood [26, 128] questioned these hypotheses because the mechanisms invoked as causes of the iron-silicate fractionations seemed to be contrived and physically implausible. He proposed instead that the relative abundances of iron and common lithophile elements (Mg, Si, Ca, Al) were the same in Mars, Earth, and Venus, and were similar to those in chondrites and in the sun. Differences in density between Mars, Venus, and Earth were caused by different oxidation states, a variable that was readily explicable on cosmochemical grounds. He reexamined the evidence on abundances of elements in the sun and concluded
that the precision of the existing data base did not justify the earlier conclusion that iron was strongly depleted [129].

Supporting evidence for this interpretation was soon forthcoming. A more accurate determination of the relative abundance of iron in the sun [130] showed that it was similar to chondrites. New space probe measurements provided precise values of the densities of Venus and Earth and of the martian moment-of-inertia. It was readily demonstrated that the gross physical (and chemical) properties of Venus, Earth, and Mars were explicable in terms of models based on chondritic abundances of major elements (except for differing mean oxidation states), with Mars being substantially more oxidized than Earth and Venus. More recently, It was shown that the metallic cores of the Earth and Venus probably contain considerable amounts of dissolved FeO [62]. This implies that the mean oxidation states of Venus, Earth, and Mars are more similar than had previously been supposed. Indeed, they could be essentially identical. The only planet displaying the effects of major metal/silicate fractionation is Mercury. The high density of Mercury, implying a large iron content, may be due to specialized conditions of accretion owing to its location nearest to the sun e.g. [131].

The considerations discussed above have an important bearing on the origin of the Moon. During the 1960's it seemed reasonable to many scientists to treat the Moon as representing an extreme case of the general process of iron/silicate fractionation that was believed to have occurred between the sun and the planets, and between the planets themselves. In this sense, it was regarded as an “independent planet” which, owing to special circumstances, had accreted in orbit around the Earth, or had been captured into Earth orbit. We now recognise that the Moon must have accreted in a region of the solar system between Earth, Mars and Venus, and probably close to the Earth. Currently, there is no evidence requiring iron/silicate fractionation between these planets and the sun. From this perspective, the large depletion of metallic iron in the Moon must be recognized as a truly remarkable phenomenon. It was this recognition that led Ringwood [24, 26] to reject a normal “planetary” origin for the Moon and to revive Darwins’s [132, 133] hypothesis that the material now in the Moon was derived by a special process from the Earth’s mantle after the core had formed. Analogous hypotheses were also developed during the 1960’s by Wise, Cameron and O’Keefe [134–136]. Although this notion of a special “earth-related” origin for the Moon was widely rejected by planetary scientists for many years, it has become popular in recent times.

According to the models developed by Ringwood [24, 26, 81, 116, 117], accretion of the Earth occurred on a short timescale (\(\sim 10^6\) yr) and was completed before the primordial gases of the solar nebula had been dissipated. Accretion was accompanied by the formation of a massive primitive terrestrial atmosphere comprised mainly of hydrogen gravitationally captured from the solar nebula and mixed with higher molecular weight gases produced by impact degassing of accreting planetesimals. The primitive atmosphere was coupled to the Earth’s rotation through turbulent viscosity and hydromagnetic torques and was thereby spun out into a co-rotating disk (period \(\sim 5\) hr). During the later stages of accretion of the Earth, high temperatures were produced by a combination of rapid accretion and thermal insulation by the primitive atmosphere. In these conditions, material from the mantle was evaporated into the primitive atmosphere and spun out into the disk. As the primitive atmosphere cooled and was dissipated, mainly by strong particle and UV radiation during the solar T-Tauri phase, the silicate components were precipitated to form a ring of Earth-orbiting planetesimals. Further fractionation due to volatility occurred during the precipitation process, since the more volatile components would be precipitated at relatively low temperatures, forming micrometre-sized smoke particles. These remained coupled to the escaping gases and hence were removed from the system. The Moon accreted from the ring of devolatilized Earth-orbiting planetesimals. The model is illustrated in Figure 15.

These models have been criticised on three principal grounds: (i) difficulties in dissipating the massive primitive terrestrial atmosphere; (ii) the short accretion timescale (\(\leq 10^6\) yr) required to provide high surface temperatures (\(>1800\) °C) on the Earth; and (iii) the problem of accounting for the high angular momentum density of the earth-moon system. Nevertheless, several of the concepts advocated in these models [24, 26] have become incorporated much later in models developed by others, some of which have achieved widespread acceptance. These include:

- Formation of a primitive terrestrial atmosphere by impact-degassing of infalling planetesimals, e.g. [137].
Fig. 15. Model showing the growth of the Earth via an accretion disk and the formation of the Moon via the ejection of material from the Earth's mantle by impacts from late-accreting planetesimals in the presence of a co-rotating primitive terrestrial atmosphere. After [93].

- Thermal blanketing of Earth by a primitive terrestrial atmosphere leading to high surface temperatures during accretion, e.g. [138].
- Redox equilibria between infalling planetesimals and the primitive atmosphere, e.g. [139].
- Dissipation of a large primitive terrestrial atmosphere by T-Tauri radiation from the early sun, e.g. [10, 140, 141].
- Accretion of the Earth in the presence of the primordial solar nebula prior to the loss of nebula gases, e.g. [10].
- Accretional heating of the Earth leading to prompt core-formation, e.g. [3].

The problem of a brief accretion timescale for the Earth in order to generate the sustained high temperatures necessary to evaporate part of the upper mantle has been short-circuited by the giant impact models of Hartmann and Davis [142] and of Cameron and Ward [143]. These authors invoke one or a few impacts by giant, late-accreting planetesimals to evaporate material from the Earth's mantle and to place it into orbit. The planetesimals are believed to have had sizes ranging between a substantial fraction of the lunar mass and the size of Mars. These models represent an important conceptual advance and provide a more plausible mechanism to evaporate terrestrial mantle material and place it into orbit than is provided by Ringwood's models. Nevertheless, as noted by Stevenson [121], the physical processes envisaged are quite closely related to those of these earlier models. Rather than achieve high mantle-evaporation temperatures via the 'steady-state' liberation of gravitational potential energy from a continuum of relatively small impacts over a short accretion timescale (~ 1 m.y.), giant impact models achieve these conditions in one or in a few transient ultrahigh energy events, and permit a much longer accretion timescale for the Earth. The impact models also lead to the formation of a circumterrestrial disk of gases and solids which evolves into a ring of earth-orbiting planetesimals that ultimately collect to form a single Moon.
(b) Constraints on Giant Impact Hypotheses

Studies of accretion processes by Hartmann and Davis [142] showed that the second largest planetesimal to impact the Earth may have been in the vicinity of 0.03 to 0.1 times the mass of the Earth and that the Earth was probably hit by many planetesimals in the size range 0.001–0.01 Earth mass (M_E). These authors pointed out that a planetesimal of about half a lunar mass impacting the Earth at 13 km/sec would have sufficient energy to eject two lunar masses of the Earth’s mantle to near-escape speeds. Hartmann and Davis hypothesized that, provided one or more of these collisions occurred subsequent to core-formation, it would have been possible to produce a geocentric ring of material derived mainly from the Earth’s mantle, from which the Moon may have then formed.

A related model was proposed by Cameron and Ward [143] who suggested that subsequent to core-formation, the Earth was hit by a mars-sized (0.1 M_E) planetesimal at about 11 km/sec. Vaporized material from the Earth’s mantle and the impactor was placed in earth-orbit, condensing to form planetesimals from which the Moon accreted. A further important property of this model is its capacity to explain the total angular momentum of the earth-moon system. The credibility of the Cameron-Ward model was further enhanced by Wetherill’s studies of planetary accretion [1, 144]. Wetherill concluded that the Earth had probably been struck by a planetesimal of martian (or larger) size. The emergence of a single mechanism with the capacity to explain both the origin of the Moon and the angular momentum density of the Earth-Moon system proved to be enormously seductive to planetary scientists. The view that the Moon had formed via impact on the Earth of a mars-sized planetesimal achieved bandwagon status, as indicated in the proceedings of the Kona conference on the Origin of the Moon [145] and in numerous subsequent papers.

Enthusiasm for this hypothesis may be premature, however. As noted in Sect. 2, it is widely believed that impact by a single mars-sized planetesimal would lead to complete or extensive melting of the Earth’s mantle. The evidence discussed in Section 2 strongly indicates that this did not occur. Accordingly the occurrence of a giant impact of this magnitude must be seriously questioned.

The hypothesis of a martian-sized impactor also encounters some additional problems. According to Wetherill’s [1, 146] accretion model, about 60% of the Earth accreted within 10 m.y. and the accretion rate declined with a half-life of ~17 m.y. Thus about 98% of the Earth would have accreted in 100 m.y. After this interval, the rate of accretion of the remaining 2% of planetesimals declined sharply. Although the accretion timescale of the Earth on this model is taken to be 100 m.y., the fate of the remaining 2% of unaccreted planetesimals deserves further consideration.

Iridium contents of lunar highland breccias show that the uppermost layer of the lunar crust contains about 1–2% of “meteoritic” (= planetesimal) material. The concentration of this component doubtless decreases with depth. If we make the generous assumption that the entire crust contains 0.5% of planetesimal component, it follows that the amount of this material which fell on the Moon after formation of the lunar crust was about 0.05% of a lunar mass. (The low rhenium abundances of mare basalts indicate that the lunar mantle contains only a minute amount, less than 0.005%, of a primary planetesimal component). Differentiation of the lunar crust seems to have been completed by about 4.45 b.y. [95, 147, 148]. At this stage, according to Wetherill’s model, about 2% of planetesimals still remained to be accreted by the terrestrial planets. This is about 40 times higher than the late planetesimal veneer which was actually accreted by the Moon. It seems doubtful whether this discrepancy can be explained adequately by a much higher capture cross section of the Earth for late planetesimals as compared to the Moon. More probably, it points to a weakness in the accretion model**.

Boss and Peale [149] estimated that the probability that an impact by a martian-sized planetesimal would produce the observed angular momentum and obliquity of the Earth-Moon system as about 1 in 20. This probability is further reduced by the remarkable timing required of this giant impact. Wetherill’s models show that giant impacts are more likely to occur at an early or intermediate stage of accretion of the Earth.

** Wetherill (pers. comm) suggests that the discrepancy is caused by the relatively high velocities of late-impacting planetesimals whose net effect is to remove surface material from the Moon. However, one would expect a substantial proportion of the ejected material to be swept up again by the Moon which would have been close to the Earth at this stage and therefore likely to experience the effects of gravitational focussing of projectile orbits caused by the Earth. The larger geometric cross section per unit mass of the Moon may have enhanced capture of this material. Clearly, the issue requires further consideration.
However, the small contents of planetesimal-derived siderophiles in the lunar crust and lunar upper mantle indicate that if formation of the Moon had been caused by a giant impact, this must have occurred at a very late stage in the accretion of the Earth. It seems that a fortuitous combination of dynamical circumstances is required to produce the Moon via the scenarios proposed in currently popular versions of the single giant impact hypothesis.

Further numerical modelling of the consequences of a martian-sized impact \cite{4, 5} have exposed another serious weakness. These workers conclude that the material which was placed in earth-orbit was derived mainly from the mantle of the impactor, rather than from the Earth’s mantle. This conclusion is at odds with the siderophile geochemistry of the Moon which points to a uniquely terrestrial signature (Figure 4), as discussed in Sects. 3 and 4. It should be a matter of some concern that the only known example of a martian-sized impactor, the SPB parent body which is presumed to be the planet Mars, possesses a siderophile signature quite unlike the Earth or the Moon (Figure 4). The difficulties of explaining the depletions of Cr, V and Mn in the Moon if it were derived from the mantle of a martian-sized planetesimal seem especially forbidding. Experimental measurements of relevant partition coefficients show that these depletions cannot be caused by core-formation within a planetesimal of this size \cite{62}.

\textit{(c) The Angular Momentum “Problem”}

In a comprehensive review of dynamical constraints on the origin of the Moon, Boss and Peale \cite{149} were deeply concerned by the “angular momentum problem”. All hypotheses of lunar origin except the giant impact model were rejected because of their perceived inability to account for the angular momentum density of the Earth-Moon system. This is unfortunate since the geochemical evidence considered in Sects. 4 and 6b oblige us also to reject the giant impact hypothesis in its present form.

Obviously, an acceptable model must satisfy the angular momentum constraint. If the Moon were extracted from the Earth’s mantle, the primordial Earth must have rotated with a period between 4 and 5 hours. Can such a short period be accommodated within the framework of acceptable models of planetary accretion? Planetary spins arise from a combination of stochastic processes (eg. large impacts) and systematic processes which can convert orbital angular momentum into prograde rotational angular momentum. Giuli \cite{150} and Harris and Ward \cite{151} showed that accretion of planetesimals possessing small eccentricities (eg. 0.01 - 0.03) systematically caused prograde planetary spins. However, accretion of planetesimals possessing larger eccentricities were as likely to contribute retrograde as prograde angular momentum. Boss and Peale \cite{149} conceded that if the Earth had accreted predominantly from planetesimals possessing very small eccentricities, a short rotation period could be provided. However, they regarded it as dynamically implausible for the Earth to accrete from a planetesimal population of this type.

Herbert et al. \cite{152} have examined a model for the formation of circumterrestrial disk and the formation of the Moon in geocentric orbit. They found that collisions between planetesimals within the Earth’s sphere of influence may cause a steady-state population of planetesimals to be captured into prograde orbits at large geocentric distances, forming an accretion disk. Owing to loss of energy via collisions within the disk, planetesimals spiral inwards and accrete upon the Earth, contributing prograde spin. Their model suggests that under favourable conditions, the total angular momentum of the Earth-Moon system could be provided.

The above authors have based their studies on dynamical models which assume that accretion of the Earth occurred after the gases of the solar nebula had been dissipated. The situation may be very different if the Earth accreted in the presence of nebula gases, e.g. \cite{10, 26}. Under these conditions, the Earth would be surrounded by a large primitive atmosphere captured from the nebula and extending out to the Hill radius \cite{10}. The inner regions of this atmosphere extending out to the geosynchronous radius may be coupled to the Earth by turbulent viscosity and may therefore co-rotate. It is possible that hydromagnetic coupling between the Earth and atmosphere may extend the co-rotation boundary further outwards than the geosynchronous radius (~ 2.5 Earth-radii for a period of 4.5 hrs). This would require early segregation of the Earth’s core to produce a magnetic field and substantial ionization of the inner regions of the terrestrial atmosphere, caused perhaps by infall of small dust particles or by T-Tauri particle and UV radiation. A sketch of this model is given in Figure 15.

In the presence of the solar nebula, aerodynamic friction will circularize the orbits of small (eg. < 10^{10} g)
planetesimals and cause them to contract radially inwards [153] until they encounter the growing planetary nucleus. If a large proportion of the mass of planetesimal population consisted of such small bodies, [11, 13], their capture by the growing Earth would have contributed efficiently towards its prograde angular momentum. Ward [154] has suggested that the orbits of much larger planetesimals (eg. \(10^{25} \text{ g}\)) may also be circularized and contract inward because of dissipation caused by density waves generated by the planetesimals in the gases of the nebula. These factors would provide the conditions required for maximizing the prograde angular momentum received by growing planets [150].

In Giuli’s gas-free accretion model [150], only those planetesimals on direct earth-encounter orbits contribute to the spin. The situation may be quite different in the presence of the nebula. As noted earlier, the Earth would capture a primitive atmosphere within its Hill sphere, and the innermost region of this atmosphere extending out to 2 or 3 earth-radii may co-rotate with the Earth. Nakazawa et al. [155] showed that low-eccentricity planetesimals entering the Hill sphere would be captured by the Earth in prograde orbits. In the absence of nebula gases, these planetesimals would orbit the Earth temporarily for a substantial number of revolutions (e.g. 10 to 100) and then escape. However, in the presence of nebula gases, the effects of gas friction may cause permanent capture on direct orbits. Nakazawa et al. [155] sought to provide a mechanism for capturing the Moon in this manner and did not pursue the implications of their model for planetary rotation. Their lunar origin model is unattractive, for several reasons, but the latter aspect may be of considerable significance. Captured planetesimals would spiral inwards owing to gas friction within the Hill sphere. Subsequently, the would transfer angular momentum to the inner co-rotating region of the atmosphere before finally encountering the Earth (Figure 15). This may provide a more effective means of delivering prograde rotational angular momentum to the Earth than impacts in a gas-free environment as in Giuli’s model. The presence of the gaseous nebula would also accelerate the sweeping up of small planetesimals, which is required in order to explain the low levels of contamination of the lunar crust by meteoritic material (Section 6b).

Further detailed investigations of these effects will be needed before it can be determined whether they can produce a primitive Earth spinning with a period of 4–5 hours. The crucial factor is the proportion of the mass of the planetesimal swarm which enters the Earth’s sphere of influence in low-eccentricity orbits. If this is large enough, rapid initial spin can be imparted.

It is well established that a significant proportion of planetesimals also encountered the Earth from orbits of relatively high eccentricity. Impacts of this type are necessary to account for the Earth’s obliquity and eccentricity. Safronov [8] estimated the minimum size of the body necessary to produce the observed obliquity as 0.001 \(M_E\). It could also have been much larger [151]. Angular momentum contributed by such impacts is necessarily stochastic in its nature.

\(d\) Lunar Origin

An attractive feature of impact models as a means of placing material from the Earth into geocentric orbit is that they rely on processes that must inevitably occur during the accretion of planets and can supply the very large amounts of energy required. It should be appreciated that present investigations of impact scenarios are at a very early stage of development and that much further work is required on the properties and sources of the impactors and the relevant impact mechanics. A perceptive discussion of these issues has been presented by Stevenson [121].

A successful model for the origin of the Moon should preferably satisfy the following conditions:

(i) Most of the material now in the Moon was derived mainly from the Earth’s mantle subsequent to the main phase of core-formation.

(ii) The impact or impacts responsible for ejection of terrestrial material did not cause complete or extensive (eg. > 50%) melting of the Earth’s mantle.

(iii) These impacts occurred before the gases of the solar nebula had been dissipated.

(iv) The impact (or impacts) occurred at a very late stage of accretion of the Earth after more than 99% of the planetesimal mass in the Earth’s neighbourhood had been swept up.

(v) The impact (or impacts) nevertheless occurred before the process of transfer of FeO from the Earth’s mantle into the core had been completed.

(vi) The proportion of vapourized material that was permanently removed from the Earth–Moon system should represent only a limited fraction (eg. < 20%) of the lunar mass, otherwise excessive chemical and isotopic fractionation of oxygen would be expected.
For reasons given previously, conditions (i), (ii) and (iv) above are not readily satisfied by models in which the Earth was struck by a mars-sized impactor [4, 5, 156]. A much wider range of possible impact scenarios has been considered by Hartmann and Vail [157], and Hartmann [158] and their results help to place some of these issues in perspective. They consider two classes of impactors: (A) silicate-iron planetesimals originating from the Earth's "feeding-zone" and possessing low approach velocities (< 5 km/sec); and (B) planetesimals composed largely of H$_2$O and other ices derived from the outer solar system and perturbed into earth-crossing orbits by Jupiter [see also, 159]. This class of planetesimals may possess approach velocities in the vicinity of 30 km/sec.

Formation of the Moon by impact(s) on the Earth from class-A planetesimals requires that they should be much smaller than Mars if melting of the terrestrial mantle is to be avoided. Impacts of this type are unlikely to spin the Earth up to a 4 hour period; hence the high angular-momentum density of the earth-moon system in this case must be provided by the non-stochastic accretion mechanisms discussed in Section 6c. It is possible that a single impact from a class-A planetesimal amounting to 0.01–0.03 M$_{\text{E}}$ might be capable of ejecting sufficient material from the Earth's mantle to form the Moon. However, quantitative studies of collisions of this kind have not yet been performed. Alternatively, multiple impacts from smaller late-accreting planetesimals, perhaps 0.001 to 0.01 M$_{\text{E}}$, might provide the necessary energy source [59, 93, 121].

Boslough and Ahrens [160] showed that planetesimals striking the Earth at 15 km/sec would vapourize ~ 5 times their own mass of target material and would shock-melt 100 times their own mass. The shock-melted material would probably have formed a spray of devolatilized droplets which may have been accelerated to high velocities by rapid expansion of the impact cloud [121, 142]. As the cloud expanded and cooled, selective condensation of the less volatile components would have produced more liquid droplets which subsequently solidified. Highly volatile elements condensed only at relatively low temperatures, forming smoke particles.

The co-rotating primitive atmosphere is believed to have played an essential role in transferring angular momentum from the solid Earth to impact-evaporated gases and liquid spray, so that a substantial proportion of the ejected material was placed in circular, equatorial orbits (Figure 15). The devolatilized droplets (1–10 mm) and coarser condensates accreted to produce an assemblage of devolatilized planetesimals. Most planetesimals probably accreted in orbits smaller than the geosynchronous limit. They would have lost energy through dissipation in the co-rotating atmosphere and hence spiralled back to Earth. However, a significant proportion of planetesimals may have been formed beyond the geosynchronous limit and beyond the Roche Limit. Here they would have accreted into moonlets. The Moon in turn is believed to have formed by cogulation from this population of moonlets.

It seems rather doubtful whether the entire Moon formed exclusively from material ejected in a single large impact. Once a proto-Moon had been formed from a single impact it seems inevitable that it would have accreted material ejected from the Earth by subsequent large impacts. The presence in lunar highlands breccias of a "primary component" bearing a strong terrestrial geochemical signature was discussed in Section 5. This component was believed to have been derived from within the Moon. The possibility should not be overlooked, however, that the primary component in the Apollo 16 breccias represents material ejected directly from the Earth's surface by impact processes [76].

Hartmann and Vail [157] also considered the possibility that the Moon was formed by the impact of an icy planetesimal from the outer solar system at a velocity of about 30 km/sec. They found that an impact by such a planetesimal amounting to 0.02–0.03 M$_{\text{E}}$ would be capable of spinning up the Earth to a period of 4–5 hours and explaining its obliquity. Whether such an impact would melt the mantle is not known. It is expected that the large proportion of rapidly expanding gases involved in an impact of this type would be highly effective in ejecting terrestrial material into earth-orbit. If the constraint that the impactor should provide most of the angular momentum of the earth-moon system is relaxed (Sect. 6c), it is possible that the impactor could have been much smaller than 0.01 M$_{\text{E}}$.

An icy impactor derived from the outer solar system would probably satisfy the criterion that the moon-forming impact occurred near the very end of the accretion of the Earth. The collision mechanics should preferably minimize exchange of oxygen between the vapourized ices of the planetesimal and shock-melted silicates from the Earth's mantle in order to preserve
the similarity of oxygen isotopic compositions. Detailed studies of the effects of high velocity impacts by icy planetesimals upon the Earth are urgently required.

7. Conclusions

In his review of dynamical constraints on the origin of the Moon presented at the 1984 Kona conference, Peale remarked that because of the diverse viewpoints on lunar origin expressed by geochemists at the conference, the geochemical evidence should be considered to cancel itself out, leaving dynamical evidence as the supreme arbiter. This view seemed to be widely shared by dynamicists present at the conference.

In the present paper, I have tried to show that the dynamical evidence may not justify the narrow interpretations that are currently being placed upon it, and that further consideration should be given to the complexities that arise when planetary accretion occurs before the gases of the solar nebula were dissipated. The geochemical evidence bearing on lunar origin is complex and diverse and requires some dedication to comprehend. This has hindered its acceptance, even among geochemists. Many of the controversies in the area have centered on entirely trivial issues. It is essential that this evidence be considered in its entirety and not on a piecemeal basis.

I believe that the corpus of evidence based upon the siderophile geochemistry of planetary mantles now constitutes a compelling case for the terrestrial origin of the Moon. This case was recognized independently and developed in parallel by the Mainz and Canberra groups. However, as is made clear in the present paper, it has been Heinrich Wänke and his colleagues who have made the major contribution, both in the acquisition of the primary data and in its interpretation. I am very pleased to have had the opportunity to highlight the importance of his achievements in this field.

Appendix

Comments Upon Recent Papers by Newsom, Taylor and Drake Dealing with Siderophile Geochemistry and Lunar Origin

The bearing of siderophile element abundances upon the origin of the Moon has been discussed by Drake, Newsom and Taylor [101, 102, 69, 162], who reach very different conclusions from those obtained in this paper. It is therefore appropriate to examine the reasons for these differences. The issues are not substantially concerned with the data base for siderophile elements in the Moon and Earth’s mantle, about which there is a reasonable level of agreement. Rather, they relate to differing interpretations of the significance of this data base.

Newsom and Drake [101] considered that the similarity of W abundances between the Earth’s mantle and Moon could be coincidental and suggested that the W content of the lunar mantle had been established by the segregation within the Moon of a geophysically plausible metallic core amounting to ~2% of the lunar mass. In order to obtain the required large depletion of W, they proposed a scenario in which segregation of metal into the core occurred at a stage when the lunar mantle had experienced only a small degree of partial melting. For a lunar mantle containing ~13% FeO, as advocated by these authors [102, 69], the degree of partial melting during core segregation is required to be in the vicinity of 1%. The physical plausibility of segregating metal under these conditions appears doubtful. They further proposed [102] a similar hypothesis to explain the P content of the Moon, but Newsom [69] later acknowledged difficulties in the model and suggested that part of the P depletion in the lunar mantle had been caused by volatilization. Newsom [69] was unable to explain the lunar geochemistry of Co and Ni in terms of the model defined above and was then obliged to increase the size of the lunar core to 5.5 wt%, maintaining that this was consistent with available geophysical constraints. However, geophysical modelling [72] shows that a core of this size lies at the extremity of the permissible range and is probably excluded if the lunar mantle contains ~13% FeO, as argued by Newsom and Drake. Moreover, a core of this size, possessing a radius of about 500 km, would be in conflict with the conclusion of Nakamura et al. [114] that normal (mantle-type) P-wave transmission occurs to a depth of 1400 km in the Moon.

Ringwood and Seifert [73] concluded that the metal phase segregating into the core from the least fractionated mare basalt source regions would contain ~40% Ni and pointed out that this would strongly decrease the metal/silicate partition coefficients employed by Newsom and Drake, thereby invalidating their models. Newsom [163] and Newsom and Taylor [162] con-
eced that their arguments would become largely irrelevant if the lunar core contained more than 15% Ni and sought to discredit Ringwood and Seifert’s conclusion. They state that the Green Glass magma which is of particular petrogenetic significance because of its high MgO content, lack of near-surface fractionation and widespread distribution at the lunar surface [79] was mistakenly assumed to represent a melt from the “primitive lunar mantle”. In actual fact, Ringwood and Seifert [73, pp. 252–254] extensively discussed the petrogenetic significance of the Green Glass magma, specifically pointing out that it was derived not from primitive mantle but from a hybridized source region composed of olivine-rich lunar mantle cumulates. This does not prevent important conclusions being drawn about the composition of those cumulates. It was shown experimentally that olivine in the Green Glass source region contained \(~830 \text{ ppm nickel (as NiO)}\) and that any metal phase in equilibrium with this olivine would therefore contain \(~40\% \text{ Ni}\) [73, 74].

It should be noted that estimates of the Ni abundance in the lunar mantle are not dependent solely upon Green Glass compositions, but have been confirmed quite independently from the well-defined NiO versus MgO abundance systematics displayed by the entire family of lunar volcanic glasses and low-Ti mare basalts [79]. It should also be recalled that the metal phase crystallizing near the liquidus of the least-fractionated low-Ti mare basalts contains 30–50% Ni [164, 165].

Newsom [163] provides an hypothetical “example” which purports to show that Green Glass (188 ppm Ni) could have been derived from a lunar mantle which itself possessed a bulk composition containing 188 ppm Ni. The argument is essentially circular and ignores most of what is known about the petrogenesis of lunar volcanic glasses in general and of Green Glass in particular. It makes unsubstantiated propositions about degrees of melting involved in forming the lunar magma ocean; moreover, its calculations are invalid because they ignore the relative proportions and types of minerals present in the lunar mantle prior to and after the melting event.

Newsom and Taylor [162] state that the oxygen fugacity of the lunar interior is far too low to permit equilibrium with a Ni-rich metallic phase. They appeal to the highly reduced state of some lunar basalts in support of their argument. However, there is strong observational evidence that this is caused largely by auto-reduction arising from loss of volatile species during crystallization of mare basalts in the near-surface environment, e.g. [164]. Detailed studies of low-Ti mare basalts have shown that their liquidus olivines contain up to 400 ppm Ni (as NiO) and that the earliest metal to crystallize contains 30–50% Ni [164, 165]. These phases, crystallizing near the liquidus of rapidly-cooled mare basalts which have experienced the smallest degrees of near-surface fractionation, provide unequivocal evidence on the high Ni contents of metal which would have been in equilibrium with their source regions. Similar evidence [74] is provided by the green volcanic glasses found at the Apollo 11, 14, 15 and 17 sites [79], which contain 130–190 ppm Ni (as NiO).

Newsom and Taylor [162] note that Cr, V and Mn are depleted (relative to CI chondrites, Mg-normalized) to similar extents in the Moon and Earth’s mantle, and attribute these depletions to loss by volatilization in the solar nebula. However, it was shown in Section 4e that Cr and V could not be depleted by this mechanism since their volatilities are respectively similar to or smaller than that of Mg. Ringwood et al. [166] have recently shown that Cr and V become siderophile under the high P, T conditions which are reached in the Earth’s lower mantle and are therefore likely to have been partitioned into the core. Their depletions in the Earth’s mantle and Moon can readily be interpreted in terms of this mechanism in the context of formation of the Moon from the Earth’s mantle.

Newsom, Taylor and Drake [101, 102, 69, 162] have proposed a series of ad hoc mechanisms to explain similarities in the abundances of certain key siderophile elements in the Earth’s mantle and Moon. We have seen above that some of these mechanisms are untenable whilst others are merely implausible. There is no reason to expect that their operation would lead to similar siderophile abundances in both bodies. To the contrary, there is very reason to expect that large differences would be produced.


[42] H. Wänke, G. Dreibus, and E. Jagoutz, Mantle chemistry and accretion history of the Earth. In „Archean


[76] See Ref. [50].


A. E. Ringwood - The Earth-Moon Connection


