GIANT IMPACT:

ACCRETION AND EVOLUTION OF THE MOON

IMPLICATIONS FOR EARTH, MARS AND THE SOLAR SYSTEM AS A WHOLE [1]

▮ J. Geiss[1], M.C.E. Huber [2] and A.P. Rossi [3] – DOI: 10.1051/epn/2014403
▮ [1] International Space Science Institute, Bern
▮ [2] Jungfraujoch Commission – Swiss Academy of Sciences
▮ [3] Jacobs University Bremen

Our planetary system has not always been as serene as it appears to us today. Exploration of the Moon has shown that disastrous collisions and violent epochs have occurred in the early part of its history. Indeed, a collision of the Earth with another planet – the Giant Impact – is the most widely accepted theory for the origin of the Moon. Several hundred million years later, Moon and Earth received a *Late Heavy Bombardment* that created the large basins on the Moon and must have devastated the atmosphere and hydrosphere of the Earth.

Article available at http://www.europhysicsnews.org or http://dx.doi.org/10.1051/epn/2014403

The exact place and time of such disastrous events are impossible to accurately retrace in time, but that does not mean they are unlikely. It has been proposed, for example, that the high density of Mercury is due to a giant impact as well, or that Uranus and Neptune migrated outwards, ravaging small-object populations until reaching their final positions several hundred million years after their formation. Also the existence of extra-solar "*Hot Jupiters*" – large planets circling close to a star – is explained by planet migration.

Obviously, the chance for life to form and survive on Earth, on Mars or on exo-planets is affected by the intensity of such violent epochs in their history.

The Giant Impact

The Giant Impact (depicted in Fig. 1) is consistent with major lunar observations, namely (i) the Moon's exceptionally large size relative to its host planet; (ii) the high angular momentum of the Earth-Moon system; (iii) the Moon's extreme depletion of volatile elements; and (iv) a differentiated global crust and mantle of the Moon that had quickly followed its delayed formation.

For the most part, the latter two findings are based on chemical analyses and age determinations of lunar rocks that had been brought back to Earth by the Apollo astronauts between 1969 and 1972.

The Giant Impact set the initial conditions for the formation and evolution of the Moon. The impactor – sometimes called Theia – is assumed to have had at least the size of Mars and to have hit the Earth at an oblique angle. The collision produced a dense protolunar cloud. Fast accretion of the Moon from this cloud assured an effective storage of gravitational energy as heat, producing early melting. A *Magma Ocean* of global dimensions formed, and upon cooling, solidified into a crust[2] and a mantle[3] (*cf.* left-hand part of Fig. 2).

The Giant Impact occurred between 70 Ma and 110 Ma after the solar-system matter had become isolated from galactic nucleosynthesis [1]. At that time, heat and accretion were the only energy sources capable of providing the energy for melting and differentiation of a lunar-sized body[4]. Heating occurs near its surface and thus much heat

[1] Editor's Note: This special Feature is an updated summary of a paper in which J. Geiss and A.P. Rossi [*The Astronomy and Astrophysics Review* (2013) On the chronology of lunar origin and evolution – Implications for Earth, Mars and the Solar System as a whole, **21**:68 (54 pp)] covered the development of lunar science over the past half-century and underlined the importance of understanding the Moon's history for the history of the Solar System as a whole. Detailed physical arguments, models, numerical estimates and references can be found in the review paper itself. *Jo Hermans, EPN Science Editor*

[2] The lunar crust is feldspar-rich, composed mainly of anorthite $CaAl_2Si_2O_8$, with a density of $\rho = 2.67$ g cm^{-3}.

[3] The mantle is composed mainly of olivines and pyroxenes $(Mg,Fe)SiO_3$ and $(Mg,Fe)_2SiO$ etc, with densities of $\rho \approx 3.5$ g cm^{-3} and $\rho \approx 3.6$ g cm^{-3}, respectively.

[4] Decay of ^{26}Al with a half-life of 0.7 Ma was the principal energy source for melting and differentiating asteroids very early in Solar-System history. At the time of the formation of the Moon, however, ^{26}Al had disappeared and the long-lived isotopes of K, U and Th with half-lives between 0.7 Ga and 14 Ga are much too slow to produce the fast melting that is inferred from the ages of lunar crust material.

is lost by radiation with increasing temperature unless a very fast accretion "buries" the heat as it is produced. Accretion times as short as $\approx 10^3$ a are needed, and model calculations show that this is achieved at the high density existing in the protolunar cloud that had been created as a consequence of the Giant Impact.

During the formation of the Moon a poorly defined intermediate layer rich in incompatible elements called KREEP developed between crust and mantle[5].

The maria: basins being filled by lava

Several 100 Ma after the Moon's accretion, the decay of the long-lived isotopes (^{40}K, ^{232}Th, ^{235}U and ^{238}U) that were concentrated in the KREEP-rich layer below the crust had generated enough internal heat to induce partial melting in the mantle. Lava extruded into large basins – *i.e.*, into the depressions that had been excavated on the Moon as part of the Late Heavy Bombardment. Upon cooling, the lava solidified into titanium-rich mare basalt (*cf.* right-hand part of Fig. 2). This era of rock formation may actually have lasted for nearly 3 Ga, from about 3.9 Ga before the present until *ca.* 1 Ga ago.

Mare basalt surfaces and other geologic units have been identified on the lunar near side by use of chemistry data from lunar orbit (*i.e.*, from remote sensing in spectral domains ranging from γ- and X-rays to the infrared). How a relative timescale was established through crater counting, and how this timescale was then calibrated by radiometric dating of rocks returned from six Apollo landing regions and three Luna[6] landing spots is presented in Box B. Detailed results, including geographical information are shown in Fig. 3.

The general, coherent explanation of the lunar origin and evolution sketched above is by and large complete. There remain some questions that still need to be resolved. In the following we briefly mention two such issues. The first one shows the precision of some lunar measurements, the second one also sheds light on early technological accomplishments.

Constraints from isotopes: the "Oxygen Isotope Puzzle"

The abundances of the chemical elements and their isotopes – as measured in the returned lunar samples – are generally compatible with the fractionation processes that go along with, and follow the Giant Impact origin of the Moon. But there remains an odd agreement of isotope ratios, the so-called *Oxygen Isotope Puzzle*, which might not *a priori* be expected as a consequence of that scenario.

The "intrinsic" isotopic composition of an object is commonly represented as Δ^{17}O, the excess – or deficit – of ^{17}O, relative to its terrestrial abundance[7]. For solar-system objects, $|\Delta^{17}O|$ values are very small, typically 1‰ or less. For Mars, for example, $\Delta^{17}O = 0.3$ ‰ has been determined[8]. Surprisingly, however, there is no significant difference in the oxygen isotopic composition between Moon and Earth.

Models have been presented indicating that this puzzle is probably not severe enough for falsifying the Giant Impact scenario. In any case, the need for explanation would be less severe if $|\Delta^{17}O|$ of Theia was smaller than the Martian value of 0.3 ‰. This is quite possible, since Theia, by definition, crossed the Earth orbit during the later phase of planet accretion, while Mars accreted its matter farther away from Earth. So it is quite possible that Theia's $|\Delta^{17}O|$ was very small.

---

[5] KREEP stands for potassium (K), rare earth elements (REE) and phosphorus (P), which together with uranium and thorium had accumulated between crust and mantle.

[6] Three of the USSR Luna spacecraft returned samples that they had collected at their landing spots.

[7] $\Delta^{17}O$ is defined such as to be unaffected by mass dependent chemical or physical processes.

decide what caused the dichotomy and when it happened might require the return and analysis of documented samples from the far side of the Moon[11].

We now turn to the use of the lunar timescale for other planets, in this case Mars, and finally discuss the lunar evidence for the Late Heavy Bombardment.

Towards a Chronology of Mars

The spatial density of craters (the 'crater frequency') can be used to infer information about the age of surfaces. Crater frequency ages and ages of Martian meteorites provide growing evidence for young lava flows, fluvial activity and aeolian deposition on Mars. In Fig. 4 plots of cumulative crater frequency[12] *vs.* diameter are shown for the ejecta blanket of the lunar North Ray Crater and for a young volcanic area on Mars.

The crater frequency age of 46.5 Ma for the lunar North-Ray Crater is in good agreement with its exposure age of 50 Ma. Due to "resurfacing", the cumulative crater frequency levels off at small diameters in both plots, but as expected, "resurfacing" for Martian areas is much faster than it is for lunar areas.

Establishing a relative timescale for Mars is now progressing fast: in the absence of documented samples from Mars, crater frequencies are being provisionally converted to actual ages (in Ma) by use of a theoretical Mars Production Function.

So far, calibration of crater frequency ages with radiometric ages or exposure ages of meteorites has not been achieved, but this is beginning to change. In a recent paper [5] it is concluded that the 55 km wide Martian Mojave Crater is the source of many Shergottites[8],

The lunar dichotomy – scientific and political surprise

In October 1959, only two years after the USSR had launched the first artificial satellite, Sputnik 1, into orbit, another Russian spacecraft, Luna-3, took images of the far side of the Moon – and that side turned out to be very different from the near side![9] Most notably, the typical *maria* seen on the near side (*cf.* Fig. 3) are almost absent.

Photographing the far side of the Moon was a technological feat: it meant on-board developing and scanning of the exposed film, so that the images could be sent back to Earth by telemetry[10].

Some of the suggested explanations for the dichotomy can be further tested by altimetry, and by gravimetric and chemical analysis from lunar orbit. However, to definitely

◄ FIG. 2 (left) The fast accretion following the Giant Impact led to a Magma Ocean covering the lunar surface, partly or fully. Before the magma solidified, light minerals (feldspar) moved towards the surface and the heavier minerals (pyroxene, olivine) sank towards the centre. The incompatible KREEP elements remained enriched in an intermediate layer. (right) Several 100 Ma after lunar accretion, lava that was melted in the mantle by the decay of long-lived isotopes of K, U and Th extruded into large basins and solidified as titanium-rich *mare* basalt. The ages indicated refer to the *Mare Imbrium* and to *mare* basalt samples collected at the Apollo 15 landing site (cf. Fig. 3).

---

[8] So far no samples have been returned from Mars, but the rare SNC meteorites are known to come from Mars. (SNC stands for the meteorite classes named after the place where the first examples had been found: Shergotti, Nakhla, Chassigny.)

[9] The Moon's rotation is tidally locked, *i.e.*, its orbital and rotational periods are the same; so there is a 'far side' of the Moon, which we never see from Earth.

[10] This happened long before today's ubiquitous CCDs were available!

[11] Detailed exploration of the lunar far side is not straightforward. It will most likely require telecommunication via a lunar relay satellite, since electromagnetic waves from the far side of the Moon do not reach the Earth.

[12] Note that crater frequencies are spatial frequencies (normally they have the dimension km⁻²).
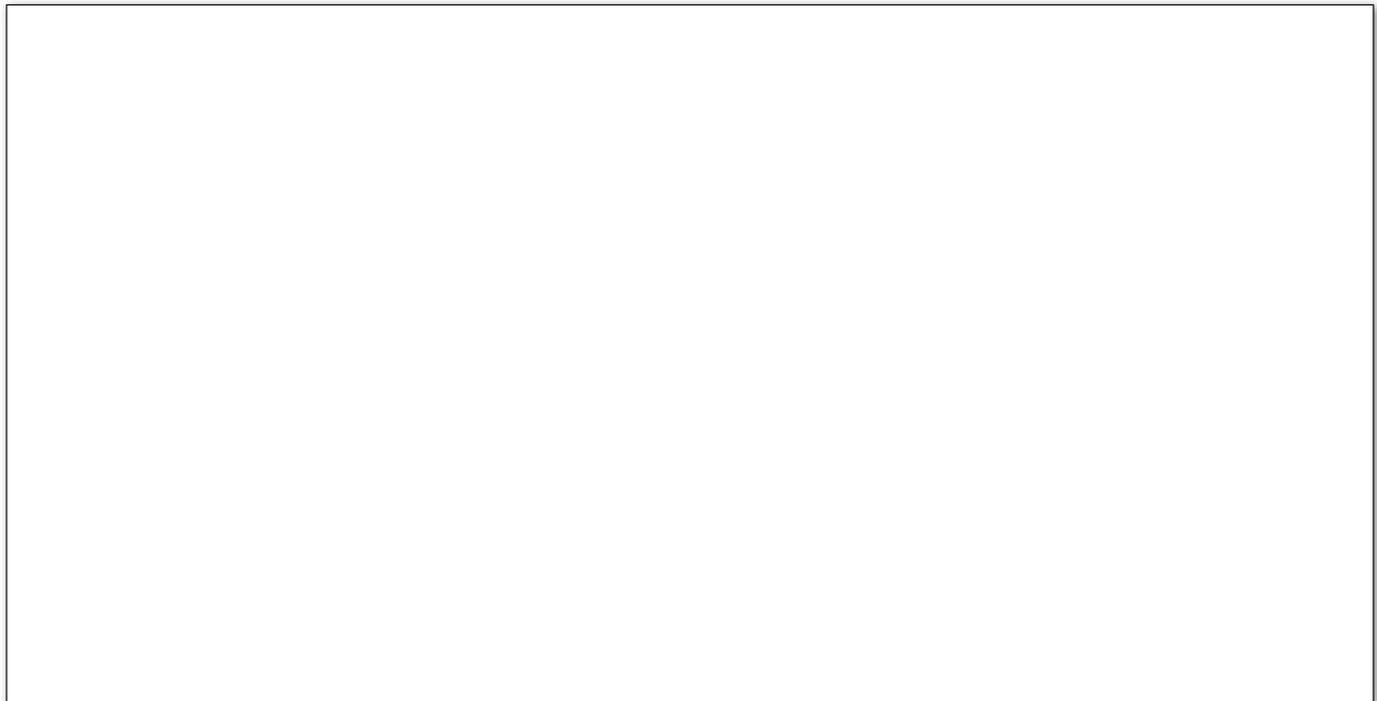
because the meteorite ejection ages coincide with the crater frequency age of that crater as calibrated with a theoretical Mars Production Function. Both ages cluster around 3 Ma. The expected production rate on Mars for Craters with a diameter D > 55 km is less than one in 10 Ma. Thus having a competitor for the Mojave Crater as source for the 3-Ma-Shergottites is rather unlikely. In fact, the ample ejection of meteorites from the young Mojave Crater might be the reason why in our epoch there are more Martian meteorites than lunar ones!

Clustering of exposure ages is observed not only for Shergottites, but also for other meteorite classes, suggesting that crater production rates are likely to vary on timescales of tens of Ma. For older surfaces such variations will be evened out when observing Cumulative Crater Frequencies.

The Late Heavy Bombardment

Radiometric dating revealed that four of the prominent large lunar basins were created during a relatively short time interval, about 4.1 Ga to 3.8 Ga before present (*cf.* Fig. 3 and the illustration in Box A). When stratigraphic observations showed that many of the other large basins on the near side of the Moon were also excavated during or close to this period, it was realised that the Moon had received a *Late Heavy Bombardment* (LHB), about 0.5 Ga to 1.6 Ga after the Moon had been accreted. Such a late clustering of major impacts must have been caused by a special phase in the history of the solar system, separate from the main accretion period. Was there a major perturbation in the planetary system so late in its history, and what could have been its cause?

Three discoveries, the lunar LHB, Trans-Neptunian objects, and Jupiter-sized Exoplanets unexpectedly close to their central stars[6] have renewed interest in planet
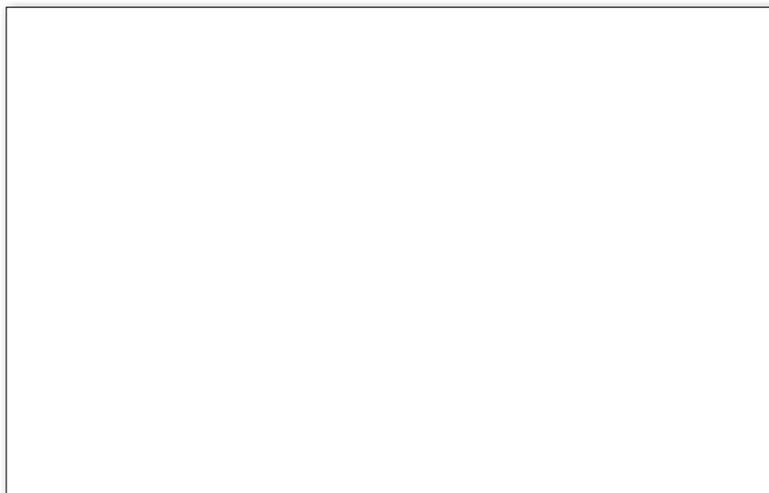
▲ FIG. 3 Two images of the near side of the Moon with chronological information. (left) Radiometric ages of four large basins (purple) and basalt samples from six *mare* areas (blue) derived from analyses of samples collected in the Apollo landing areas and the Luna landing spots (both of them marked on the right-hand image). (right) Crater frequency ages of *mare* areas from high-resolution crater counting and mapping. The youngest mare areas are in the *Oceanus Procellarum* to the West of *Mare Imbrium*; they correspond to mare areas with a high KREEP content.

◄ FIG. 4 Crater frequency ages for the lunar North Ray Crater [2] and a martian volcanic area on Mars [3] Images used were from the NASA Lunar and Mars Reconnaissance Orbiters, respectively. The Crater Frequency ages given in these plots were calculated by using the theoretical Production-Function and Chronology-Function given by Ivanov [4] (see Box B).

migration, and led to the development of Nice Models[13] for explaining late heavy bombardment epochs by planet rearrangement well after the birth of the Solar System.

The classical Nice model assumes planets to have been accreted from the proto-planetary disc more closely together than we find them today. After dissipation of the gaseous portion of the solar nebula, the planets slowly migrated, due to their interacting with the remaining planetesimals. Saturn moved away from Jupiter, increasing the ratio between the periods of the two planets (P_S and P_J), until they got into the 1:2 resonance at $P_J/P_S = 1/2$ (today, we have $P_J/P_S = 1/2.48$). The orbit of Saturn became very eccentric, causing close encounters among large planets and "a massive delivery of planetesimals to the inner solar system" that could have caused the lunar LHB.[7]

For comparisons with specific Nice-model predictions, the lunar LHB would have to be better defined than it is today. The Orientale basin is the youngest among the major lunar basins and may mark the end of the lunar LHB. Its beginning is difficult to measure and even to define. Determining instead the declining phase of the bombardment by measuring the number of basin-forming impacts in a given time interval near 4 Ga would be easier and could significantly constrain theories of LHB origin. Geiss & Rossi[1] suggest that two rover/sample-return missions would do the job, one sampling the blanket of the *Orientale* basin and the other exploring a region to the Southeast of *Mare Nectaris*[14].

The heavy bombardment of the Moon about 4.1 Ga to 3.8 Ga ago must have hit the Earth as well, causing lasting devastation of our atmosphere and hydrosphere[15]. Quantitative comparison of the effects of bombardment of Moon and Earth depends on the origin of the impactor. For objects arriving with high energy from, *e.g.*, the region beyond Uranus, impact energies on the Earth are only modestly increased, and the Moon and Earth ratio of impact rates are nearly proportional to their geometrical cross sections (1:13.4). On the other hand, for objects coming from the inner edge of the asteroid belt, focussing and acceleration by the terrestrial gravitational field strongly increases the relative flux and also the impacting energy hitting the Earth.

Remnants of basins produced on Earth during the Archean and Proterozoic epochs have been wiped out by plate tectonics. However, the distribution of spherule beds, interpreted as ejecta from large impacts, indicate

---

[13] Nice models were first developed in Nice (France).

[14] Following the delivery to the lunar surface of the Chinese rover Chang'e 3 in December 2013, a sample return mission, Chang'e 5 is scheduled for 2017.

[15] The energy of the Imbrium impact on the Moon was roughly a thousand times higher than the energy of the Chixculub impact on Earth that caused the mass extinction at the K/T boundary.

BOX A: TIMESCALES AND CHRONOLOGIES[9]

On Earth, a global relation was found in the 19th century between fossil species and sedimentary rock strata, establishing a relative chronology from the Cambrian to the present. In the 20th century this chronology was calibrated by radiometric dating and extended deep into the Precambrian. On the Moon, where there are no fossils, relative timescales are based on the observed impact crater frequencies. Crater count ages have now been determined for a large portion of the lunar surface and calibrated with radiometric and some exposure ages of lunar samples returned from six Apollo landing areas and three Luna landing spots. Even from the rather limited geographical coverage available after the Apollo and Luna missions, it was concluded that the crater production rate decreased strongly for ages larger than 3.2 Ga and remained approximately constant afterwards. By approximating the crater production rate by a constant plus an exponential, Neukum, Ivanov and Hartmann[10] introduced the Cumulative Crater Frequency Timescale shown in the illustration below; this has since been the standard. The timescale is well established only in the time interval 4 Ga to 3 Ga. It cannot be much extended beyond 4 Ga, because there the crater frequency is approaching saturation. Samples from the young areas to the West of Mare Imbrium (*cf.* Fig. 3) could close the gap in the calibrated timescale from 3 Ga to 1 Ga. Completing the lunar crater count timescale is of foremost importance, because it will remain indispensable for quite some time, as it can help for transferring absolute timescales to Mars and other solar system bodies.

▼FIG. 5 Timescale of events on the Moon, traced by the relative scale established by crater count which was calibrated by radiometry of lunar rock samples. Red, blue and purple data points represent the Late Heavy Bombardment, the *maria* and relatively young craters, respectively
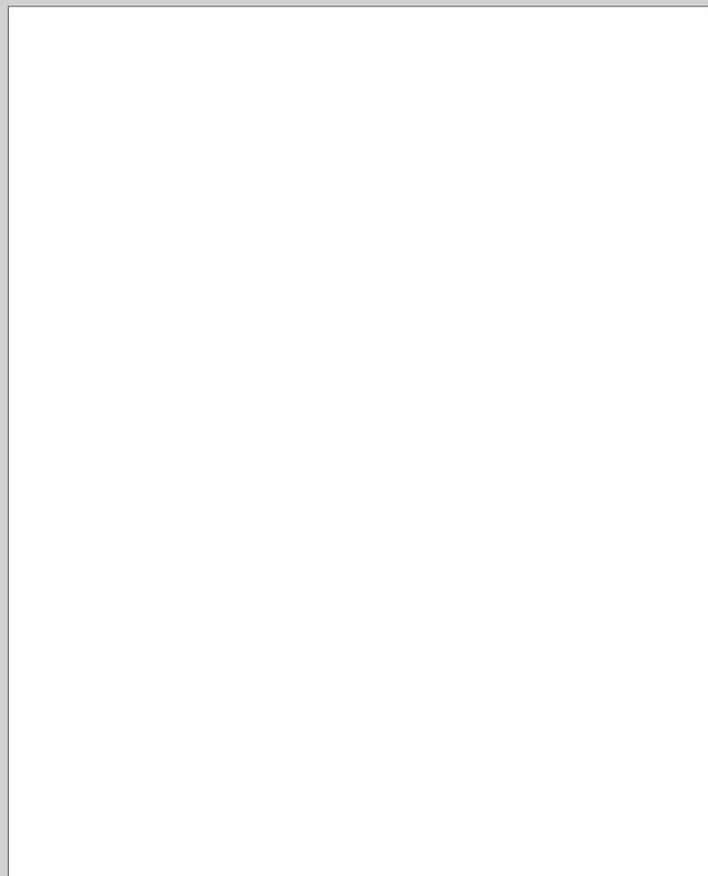
that basin forming impacts on Earth lasted to 3 Ga before the present and beyond.[8] The *lunar* LHB ended earlier, as inferred from the crater frequency curve in the illustration of Box A or from the estimated age of the Orientale basin.

The earliest, well-established traces of life appeared about 3.5 Ga ago. The analysis of spherule beds implies, however, that the epoch of heavy bombardment on Earth had not yet ended at that time. Further data on the LHB and post-LHB epochs on Moon, Earth and Mars as well as on early traces of life on Earth might tell us whether a heavy bombardment had delayed the creation of life on our planet, whether all life was extinguished that had existed before, or whether a primitive life form managed somehow to survive the epoch of heavy bombardment. ∎

BOX B: AGE DETERMINATION

Radiometric Ages are determined by radioactive dating. This technique is based on a comparison between the observed abundance of naturally occurring radioactive isotopes and their decay products, by use of the known decay rates.

Cosmic Ray Exposure Ages are based on the following model: a collision excavates a rock that was buried and screened from cosmic rays and ejects it into an orbit around the Sun, where it is fully exposed to cosmic rays. Eventually, the orbiting rock will collide with the Earth, where cosmic-ray exposure will virtually cease. The time of exposure to cosmic rays is determined from nuclides produced by cosmic-ray-induced nuclear reactions.

An exposure age (T_{EXP}) is calculated from the abundances of a pair of a stable (N_{ST}) and a radioactive (N_{RAD}) cosmic-ray-produced nuclide, the corresponding production rates P_{ST} and P_{RAD}, and the decay constant λ of the radioactive nuclide. If P_{ST} and P_{RAD} were constant over the time interval T_{EXP} to T_0, then for the simple case of $\lambda T_0 = 0$ and $\lambda T_{EXP} \gg 1$, the exposure age is given by

$$T_{EXP} = \frac{1}{\lambda} \frac{P_{RAD}}{P_{ST}} \frac{N_{ST}}{N_{RAD}}$$

This relation is valid only if the irradiation conditions (cosmic-ray intensity and energy spectrum, heliospheric modulation, shielding by atmosphere or overlying solid material) were constant from T_{EXP} to T_0. These conditions are usually met when calculating the time of ejection of a meteorite from an asteroid, Mars or Moon.

Crater Count Ages are defined as the Cumulative-Crater-Frequency [km^{-2}] = Production-Function (D) x Chronology-Function (T) [km^{-2}]. Because it is possible to separate the variables, the data can be presented in two types of plots:

• Cumulative-Crater-Frequency versus Age (see graph in BOX A) and
• Cumulative-Crater-Frequency versus D, *i.e.*, the lower integration limit for the crater diameter.

The slope of the plot (*cf.* Fig. 4) determines the age of the cratered area. This must, however, be calibrated by radiometric or exposure ages of rocks, determined in the laboratory.

About the Authors
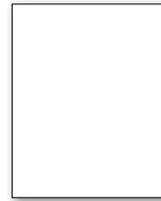
Johannes Geiss is Honorary Director and was Executive Director (1995-2002) of the International Space Science Institute in Bern. He was Professor of Physics at the University of Bern (1960-1991), Associate Professor at the Marine Laboratory of the University of Miami (1958-1959) and Research Associate at the Enrico Fermi Institute, University of Chicago (1955-1956). He is a Foreign Associate of the US National Academy of Sciences and a Member of the Academia Europaea. He received the NASA Medal for Exceptional Scientific Achievement, the Albert Einstein Medal of the Einstein Association in Bern, and the Bowie Medal of the American Geophysical Union.
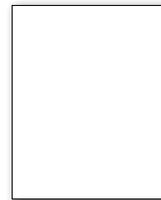
Martin C. E. Huber was Adjunct Professor at ETH Zurich, Head of ESA's Space Science Department and President of EPS (2004-2005). He was closely associated with Europhysics Letters (now EPL) and is a member of the Editorial Board of The Astronomy and Astrophysics Review. He is an Associate of Harvard College Observatory, a Member of the Academia Europaea and immediate past President of the Kommission für die Hochalpine Forschungsstation Jungfraujoch of the Swiss Academy of Sciences (SCNAT).
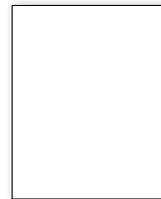
Angelo Pio Rossi is a planetary scientist based at Jacobs University Bremen. He is a geologist by background and works on Earth and Planetary Remote Sensing and Comparative Planetology. He is a Co-I of the High Resolution Stereo Camera on board of ESA's Mars Express and a member of the Editorial Boards of Planetary and Space Science and EGU Solid Earth. Previously he was a Research Fellow in ESA's Science Directorate and Staff Scientist in the International Space Science Institute.

References

[1] see A.N. Halliday, *Phil Trans R Soc A* **366**, 4163 (2008).

[2] H. Hiesinger, C.H. Van der Bogert, J.H. Pasckert *et al.*, *J Geophys Res* **117**, E00H10 (2010).

[3] F. Jagert, E. Hauber, *Geoinf* **2012 (2)**, 177 (2012).

[4] B.A. Ivanov, *Space Sci Rev* **96**, 87 (2001).

[5] S.C. Werner, A. Ody, F. Poulet, *Science* **343**, 1343 (2014).

[6] M. Mayor & D. Queloz, *Nature* **378**, 355 (1995).

[7] R. Gomes, H.F. Levison, K. Tsiganis, A. Morbidelli, *Nature* **435**, 466 (2005).

[8] B.C. Johnson and H.J. Melosh, *Nature* **485**, 75 (2012); W.F. Bottke, D. Vokrouhlicky, D. Minton *et al.*, *Nature* **485**, 78 (2012).

[9] R. Kallenbach, J. Geiss and W.K. Hartmann, *Space Sci Ser ISSI (SSSI)* **12** (2001).

[10] G. Neukum, B.A. Ivanov, W.K. Hartmann, *Space Sci Rev* **96**, 55 (2001), *and SSSI* **12**, 55 (2001).